æ

1. Introduction

Stability and sensitivity of optimization problems are understood here, respectively, as local Lipschitz continuity and differentiability of solutions of such problems with respect to parameters. Due to the presence of inequality type constraints, optimization problems are nonsmooth even for arbitrarily smooth data. That is the reason why there are fundamental difficulties in application of the classical implicit function theorem in sensitivity analysis.

The systematic studies of stability and sensitivity of solutions to optimization problems were initiated in mid-seventies with two important papers by S. M. Robinson [43] and A. V. Fiacco [18], which concerned parametric mathematical programs in finite dimensions. A fundamental break-through was made in the paper "Strongly regular generalized equations" by S. M. Robinson [44], in which the author developed an implicit function theorem for generalized equations (inclusions). The theorem is formulated and proved in Banach spaces and it provides an adequate mathematical tool for studying stability for a broad class of parametric optimization problems. In mid-nineties an important generalization of Robinson's theorem was developed by A. L. Dontchev [10], who extended the original result to nonlinear topological spaces. The original application of Robinson himself, as well as of some other authors, concerned mathematical programs in finite dimensions. However, cone-constrained optimization problems in abstract Banach spaces (see, e.g., [6, 8, 33, 34, 49]), as well as semi-infinite programs [5, 28, 29] and optimal control problems [1, 2, 13, 15, 33, 34, 35] were also investigated. For more references cf. the bibliography in [7, 8].

The approach based on Robinson's implicit function theorem does not provide any information on differentiability of solutions with respect to parameters. In sensitivity analysis, the first investigations concerned again finite-dimensional mathematical programs. In [43] and [18], the classical implicit function theorem was used, which required, among other things, the assumption of strict complementarity. Later, this last assumption was dropped and conditions of directional [25, 26] and Bouligand [44, 48] differentiability were obtained.

In infinite-dimensional optimization problems, the concept of *polyhedric sets* was introduced and applied to get directional differentiability of solutions to problems with linear constraints [21, 41]. For applications to optimal control cf. [34, 50, 51].

Robinson further developed his implicit function theorem in [46], where the concept of *strong approximation* was introduced and exploited. In particular, the theorem allows investigating Bouligand differentiability of solutions to constrained optimization problems.

Important extensions of this theorem are due to A. L. Dontchev [9], who, among other things, proved that various types of differentiability of solutions to nonlinear optimization problems are equivalent to the same type of differentiability of solutions to linear-quadratic accessory problems. Actually, the theorems due to Robinson and Dontchev allow reducing stability and sensitivity analysis for nonlinear parametric problems to the same analysis for linear-quadratic accessory problems subject to additive perturbations. These last problems are usually much easier to investigate than the original ones.

In parametric optimal control Robinson's theorem was first applied by Ito and Kunisch [25]. Further development can be found in [33, 13, 35, 15, 14]. In [15] the most refined sufficient conditions of Lipschitz stability for control constrained optimal control problems were derived.

The above approach allows one to obtain *sufficient* conditions of stability and sensitivity, but it does not provide any information on how close these conditions are to being necessary. Recent results show that for mathematical programs [17, 12] and control constrained optimal control problems [16], sufficient optimality conditions are also *necessary*, provided that the dependence of data on the parameter is *sufficiently strong*. Thus, we get a full characterization of stability and sensitivity properties.

The purpose of this paper is to present systematically the stability and sensitivity analysis for nonlinear optimal control problems subject to mixed control-state constraints. For this class of problems the analysis is more or less complete. The presented sufficient conditions of Lipschitz stability are mostly based on material known in the literature. The main elements of novelty concern differentiability of the solutions and necessary conditions of stability and sensitivity.

The organization of the paper is the following. In Sections 2 and 3 the problem is stated and the basic assumptions are introduced. The abstract theorems which are the main mathematical tools of the further analysis are recalled in Section 4. Sections 5 and 6 are devoted to stability and sensitivity analysis for the accessory linear-quadratic problem. The basic stability and sensitivity results for the original nonlinear problem are derived in Sections 7 and 8. In Section 9 the necessity of the assumptions applied is discussed. The principal result is stated in Section 10, where also some concluding remarks and open questions are formulated. The main sections end with short bibliographical notes.

We use the following notations: Capital letters $X, Y, Z, \ldots$, sometimes with superscripts, denote Banach or Hilbert spaces. The norms are denoted by $\|\cdot\|$ with a subscript referring to the space.

$$\mathcal{B}_\varrho^X(x_0) := \{x \in X \mid \|x - x_0\|_X \leq \varrho\}$$

is the closed ball in X of radius ϱ, centered at x_0. Asterisks denote dual spaces, as well as dual operators. (y, x), with $x \in X$ and $y \in X^*$, is the duality pairing between X and X^*. The inner product in a Hilbert space X is denoted by $(\cdot, \cdot)_X$.

For $f : X \times Y \to Z$ let $D_x f(x, y), D_y f(x, y), D^2_{xy} f(x, y), \ldots$ denote the respective Fréchet derivatives in the corresponding arguments.

$\mathbb{R}^n$ is the n-dimensional Euclidean space with inner product $\langle x, y\rangle$ and norm $|x| = \langle x, x\rangle^{1/2}$. A superscript i denotes the ith component of a vector or the ith row of a matrix, while a superscript ij refers to the appropriate element of a matrix. Transposition is denoted by $*$.

For a given subset $I \subset \mathbb{R}$, $L^s(I;\mathbb{R}^n)$ is the Banach space of measurable functions $f : I \to \mathbb{R}^n$, supplied with the norm

$$\|f\|_s = \begin{cases} [\int_I |f(t)|^s \, dt]^{1/s} & \text{for } s \in [1,\infty), \\ \operatorname{ess\,sup}_{t \in I} |f(t)| & \text{for } s = \infty. \end{cases}$$

$W^{1,s}(I;\mathbb{R}^n)$ denotes the Sobolev space of functions f absolutely continuous on I with the norm

$$\|f\|_{1,s} = \begin{cases} [|f(0)|^s + \|\dot{f}\|_s^s]^{1/s} & \text{for } s \in [1,\infty), \\ \max\{|f(0)|, \|\dot{f}\|_\infty\} & \text{for } s = \infty. \end{cases}$$

c, k and l denote generic constants, not necessarily the same in different places.

2. Problem formulation and constraint qualifications

Let us introduce the space $H = \mathbb{R}^r \times L^\infty(0,1;\mathbb{R}^r)$ of parameters. Let $G \subset \mathbb{R}^r$ be an open bounded set of feasible values of parameters and let

$$\mathcal{G} = \{h \in H \mid h(0) \in G \text{ and } h(t) \in G \text{ for a.a. } t \in [0,1]\}$$

denote the open set of feasible parameters. Consider the family of the following optimal control problems depending on $h \in \mathcal{G}$:

(O)$_h$ Find $(x_h, u_h) \in X^\infty$ such that

$$F(x_h, u_h, h) = \min\Big\{F(x,u,h) := \int_0^1 \varphi(x(t), u(t), h(t))\, dt + \psi(x(0), x(1), h(0))\Big\} \tag{2.1}$$

subject to

$$\dot{x}(t) - f(x(t), u(t), h(t)) = 0 \quad \text{for a.a. } t \in [0,1], \tag{2.2}$$

$$\xi(x(0), x(1), h(0)) = 0, \tag{2.3}$$

$$\theta(x(t), u(t), h(t)) \le 0 \quad \text{for a.a. } t \in [0,1], \tag{2.4}$$

where $X^s = W^{1,s}(0,1;\mathbb{R}^n) \times L^s(0,1;\mathbb{R}^m)$, $s \in [1,\infty]$,

$$\varphi : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^r \to \mathbb{R}, \qquad \psi : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R},$$
$$f : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^r \to \mathbb{R}^n, \quad \xi : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}^n,$$
$$\theta : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^r \to \mathbb{R}^l.$$

The following standing assumptions are supposed to be satisfied throughout the paper.

(I) There exist open sets $\mathcal{R}^n \subset \mathbb{R}^n$ and $\mathcal{R}^m \subset \mathbb{R}^m$ such that the functions $\varphi(\cdot,\cdot,\cdot)$, $f(\cdot,\cdot,\cdot)$, $\theta(\cdot,\cdot,\cdot)$ and $\psi(\cdot,\cdot,\cdot)$, $\xi(\cdot,\cdot,\cdot)$, are twice Fréchet differentiable in (x,u) on $\mathcal{R}^n \times \mathcal{R}^m \times G$ and on $\mathcal{R}^n \times \mathcal{R}^n \times G$, respectively. All these functions, together with their first derivatives in (x,u), are Fréchet differentiable in h. All first and second order derivatives are uniformly Lipschitz continuous in (x,u,h).

(II) For a given reference value $h_0 \in \mathcal{G}$ of the parameter there exists a reference solution $(x_0, u_0) := (x_{h_0}, u_{h_0})$ of (O)$_{h_0}$ and $(x_0(t), u_0(t)) \in \mathcal{R}^n \times \mathcal{R}^m$ for a.a. $t \in [0,1]$.

To simplify notation, we will denote by subscript "0" the functions evaluated at the reference point, e.g., $\varphi_0 := \varphi(x_0, u_0, h_0)$, $\psi_0 := \psi(x_0(0), x_0(1), h_0(0))$. Moreover, we define

$$\begin{aligned} A(t) &= D_x f(x_0(t), u_0(t), h_0(t)), & B(t) &= D_u f(x_0(t), u_0(t), h_0(t)), \\ \Xi_0 &= D_{x(0)}\xi(x_0(0), x_0(1), h_0(0)), & \Xi_1 &= D_{x(1)}\xi(x_0(0), x_0(1), h_0(0)), \\ \Upsilon(t) &= D_x\theta(x_0(t), u_0(t), h_0(t)), & \Theta(t) &= D_u\theta(x_0(t), u_0(t), h_0(t)). \end{aligned} \tag{2.5}$$

Our last standing assumption is:

(III) $\operatorname{rank}[\Xi_0 + \Xi_1\phi(1)] = n$,

where Φ is the fundamental matrix solution of $\dot y - Ay = 0$:

$$\dot\Phi(t) - A(t)\Phi(t) = 0, \quad \Phi(0) = I. \tag{2.6}$$

We will need some constraint qualifications. To this end, let us denote by $I = \{1, \ldots, l\}$ the set of the indices of inequality constraints and for $\alpha \geq 0$ and $t \in [0,1]$, introduce the subsets of α-active constraints:

$$I_\alpha(t) = \{i \in I \mid \theta_0^i(t) \geq -\alpha\}, \quad \imath_\alpha(t) = \operatorname{card} I_\alpha(t). \tag{2.7}$$

Define the functions

$$\theta_\alpha^i(t) = \min\{\theta_0^i(t) + \alpha, 0\}, \tag{2.8}$$

as well as the $l \times l$ and $l \times (m+l)$-matrices

$$T_\alpha(t) = \operatorname{diag}\theta_\alpha^i(t), \quad V_\alpha(t) = [\Theta(t),\ T_\alpha(t)]. \tag{2.9}$$

We assume

(A1) (*Linear Independence*) There exist constants $\alpha, \beta > 0$ such that

$$|V_\alpha(t)V_\alpha(t)^*\eta| \geq \beta|\eta| \quad \text{for all } \eta \in \mathbb{R}^l \text{ and a.a. } t \in [0,1]. \tag{2.10}$$

(A2) (*Controllability*) There exists $\alpha > 0$ such that, for each $e \in \mathbb{R}^n$, there exist $v \in L^\infty(0,1;\mathbb{R}^m)$, $\vartheta \in L^\infty(0,1;\mathbb{R}^l)$ and $y \in W^{1,\infty}(0,1;\mathbb{R}^n)$ which satisfy the following equations:

$$\dot y - Ay - Bv = 0, \tag{2.11}$$

$$\Xi_0 y(0) + \Xi_1 y(1) = e, \tag{2.12}$$

$$\Upsilon y + \Theta v + T_\alpha\vartheta = 0. \tag{2.13}$$

Clearly, if (2.10) is satisfied for an $\overline{\alpha} > 0$, then it is also satisfied for any $\alpha \in [0, \overline{\alpha}]$. Define

$$\begin{aligned} \widehat\Theta_\alpha(t) &= [D_u\theta_0^i(t)]_{i\in I_\alpha(t)}, & \widehat\Upsilon_\alpha(t) &= [D_x\theta_0^i(t)]_{i\in I_\alpha(t)}, \\ \widehat T_\alpha(t) &= [\operatorname{diag}\theta_\alpha^i(t)]_{i\in I_\alpha(t)}, & \widehat\mu(t) &= [\mu^i(t)]_{i\in I_\alpha(t)}. \end{aligned} \tag{2.14}$$

The following lemma relates condition (A1) to the usual constraint qualification in the form of linear independence of gradients of active constraints.

LEMMA 2.1. *Condition* (A1) *holds if and only if there exist constants* $\widehat\alpha, \widehat\beta > 0$ *such that*

$$|\widehat\Theta_{\widehat\alpha}(t)^*\widehat\mu| \geq \widehat\beta|\widehat\mu| \quad \textit{for all } \widehat\mu \in \mathbb{R}^{\imath_{\widehat\alpha}(t)} \textit{ and a.a. } t \in [0,1], \tag{2.15}$$

i.e., pointwise, the gradients of $\widehat{\alpha}$*-active constraints are linearly independent, uniformly on* $[0,1]$.

Proof. Suppose that (2.15) holds. Since by (2.8),

$$|\theta_0^i(t)| \begin{cases} \le \widehat{\alpha} & \text{for } i \in I_{\widehat{\alpha}}(t), \\ > \widehat{\alpha} & \text{for } i \notin I_{\widehat{\alpha}}(t), \end{cases} \tag{2.16}$$

in view of (2.9), we have

$$|[\Theta(t),\ T_{\widehat{\alpha}}(t)]^* \mu| \ge \min\{\widehat{\beta}, \widehat{\alpha}\}|\mu| \quad \text{for all } \mu \in \mathbb{R}^n,$$

i.e., (2.10) is satisfied with $\beta = (\min\{\widehat{\beta}, \widehat{\alpha}\})^2$.

Suppose now that (2.10) holds. Choose $\widehat{\alpha} = \sqrt{\beta}/2$ and put $\mu^i = 0$ for $i \notin I_{\widehat{\alpha}}(t)$. In view of (2.16) we obtain

$$|\widehat{\Theta}_{\widehat{\alpha}}(t)^* \widehat{\mu}| \ge |[\widehat{\Theta}_{\widehat{\alpha}}(t),\ \widehat{T}_{\widehat{\alpha}}(t)]^* \widehat{\mu}| - |\widehat{T}_{\widehat{\alpha}}(t)^* \widehat{\mu}| \ge \sqrt{\beta}|\widehat{\mu}| - \frac{\sqrt{\beta}}{2}|\widehat{\mu}| = \frac{\sqrt{\beta}}{2}|\widehat{\mu}|,$$

i.e., (2.15) holds with $\widehat{\beta} = \sqrt{\beta}/2$. ■

It will be convenient to express the controllability condition (A2) in an equivalent form.

LEMMA 2.2. *If* (A1) *holds, then* (A2) *is equivalent to the condition that for any* $e \in \mathbb{R}^n$ *there exists a solution* $(y, \zeta) \in W^{1,\infty}(0,1;\mathbb{R}^n) \times L^\infty(0,1;\mathbb{R}^{m+l})$ *of the equation*

$$\dot{y} - \mathcal{A}_\alpha y - \mathcal{B}_\alpha \zeta = 0, \quad \Xi_0 y(0) + \Xi_1 y(1) = e, \tag{2.17}$$

where

$$\mathcal{A}_\alpha = A - [B,\ 0]V_\alpha^*(V_\alpha V_\alpha^*)^{-1}\Upsilon, \quad \mathcal{B}_\alpha = [B,\ 0](I - V_\alpha^*(V_\alpha V_\alpha^*)^{-1}V_\alpha). \tag{2.18}$$

Proof. In view of (A1) any solution (v, ϑ) of (2.13) can be expressed in the form

$$\begin{bmatrix} v \\ \vartheta \end{bmatrix} = -V_\alpha^*(V_\alpha V_\alpha^*)^{-1}\Upsilon y + (I - V_\alpha^*(V_\alpha V_\alpha^*)^{-1}V_\alpha)\zeta, \tag{2.19}$$

where $\zeta \in \mathbb{R}^{m+l}$ is arbitrary. Substituting (2.19) into (2.11) we obtain (2.17). ■

Define the spaces

$$\left.\begin{aligned} Z^s &= W^{1,s}(0,1;\mathbb{R}^n) \times L^s(0,1;\mathbb{R}^m) \times L^s(0,1;\mathbb{R}^l), \\ Y^s &= L^s(0,1;\mathbb{R}^n) \times \mathbb{R}^n \times L^s(0,1;\mathbb{R}^l), \\ Y^{1,s} &= W^{1,s}(0,1;\mathbb{R}^n) \times \mathbb{R}^n \times L^s(0,1;\mathbb{R}^l), \end{aligned}\right\} s \in [1,\infty],$$

and the mapping $\mathcal{C}_\alpha : Z^s \to Y^s$ given by the left-hand side of (2.11)–(2.13):

$$\mathcal{C}_\alpha \begin{pmatrix} y \\ v \\ \vartheta \end{pmatrix} = \begin{pmatrix} \dot{y} - Ay - Bv \\ \Xi_0 y(0) + \Xi_1 y(1) \\ \Upsilon y + \Theta v + T_\alpha \vartheta \end{pmatrix}. \tag{2.20}$$

LEMMA 2.3. *Assumptions* (A1) *and* (A2) *imply that* $\mathcal{C}_\alpha : Z^s \to Y^s$ *is surjective for any* $s \in [1,\infty]$. *If* $\mathcal{C}_\alpha : Z^s \to Y^s$ *is surjective for any* $s \in [1,\infty)$ *then* (A1) *and* (A2) *hold.*

Proof. First, let us prove that (A1) and (A2) imply surjectivity of $\mathcal{C}_\alpha$. We have to show that for any $(p,q,r) \in Y^s$ the equation

$$\begin{aligned} \dot{y} - Ay - Bv &= p, \\ \Xi_0 y(0) + \Xi_1 y(1) &= q, \\ \Upsilon y + \Theta v + T_\alpha \vartheta &= r, \end{aligned} \tag{2.21}$$

has a solution $(y,u,\vartheta) \in Z^s$. As in (2.19), from the last equation in (2.21) we obtain

$$\begin{bmatrix} v \\ \vartheta \end{bmatrix} = V_\alpha^*(V_\alpha V_\alpha^*)^{-1}(r - \Upsilon y) + (I - V_\alpha^*(V_\alpha V_\alpha^*)^{-1}V_\alpha)\zeta, \tag{2.22}$$

and (2.21) yields

$$\dot{y} - \mathcal{A}_\alpha y - \mathcal{B}_\alpha \zeta = \widetilde{p}, \quad \Xi_0 y(0) + \Xi_1 y(1) = q, \tag{2.23}$$

where $\mathcal{A}_\alpha$ and $\mathcal{B}_\alpha$ are given in (2.18) and

$$\widetilde{p} = p + [B,\ 0]V^*[VV^*]^{-1}r. \tag{2.24}$$

Let $\Phi_\alpha(t)$ denote the solution of the homogeneous equation for symmetric matrix functions

$$\dot{\Phi}_\alpha(t) - \mathcal{A}_\alpha(t)\Phi_\alpha(t) = 0, \quad \Phi_\alpha(0) = I.$$

Then the solution of the equation

$$\dot{x} - \mathcal{A}_\alpha x - \widetilde{p} = 0, \quad x(0) = 0,$$

is given by

$$x(t) = \int_0^t \Phi_\alpha(t)^{-1}\Phi_\alpha(\tau)\widetilde{p}(\tau)\,d\tau. \tag{2.25}$$

Let us introduce the new variable

$$z(t) = y(t) - x(t). \tag{2.26}$$

Then from (2.23) we obtain

$$\dot{z} - \mathcal{A}_\alpha z - \mathcal{B}_\alpha \zeta = 0, \quad \Xi_0 z(0) + \Xi_1 z(1) = \widetilde{q}, \tag{2.27}$$

where, in view of (2.24) and (2.25),

$$\widetilde{q} = q - \Xi_1 \int_0^1 \Phi_\alpha(1)^{-1}\Phi_\alpha(\tau)\{p(\tau) + [B(\tau),\ 0]V(\tau)^*[V(\tau)V(\tau)^*]^{-1}r(\tau)\}\,d\tau.$$

By (A1), (A2) and Lemma 2.2, equation (2.27) has a solution $(z,\zeta) \in W^{1,\infty}(0,1;\mathbb{R}^n) \times L^\infty(0,1;\mathbb{R}^{m+l})$, which, by (2.25), (2.26) and (2.22), corresponds to a solution $(y,u,\vartheta) \in Z^s$ of (2.21). This completes the proof of surjectivity of $\mathcal{C}_\alpha$.

Let us now show the opposite implication. Assume that, for some $s \in [1,\infty)$, equation (2.21) has a solution for any $(p,q,r) \in Y^s$. Setting $p = 0$ and $r = 0$, we obtain (A2). To show (A1), assume that (2.10) is violated, i.e., there exists a subset $M \subset [0,1]$ with $\operatorname{meas} M > 0$ such that $\operatorname{range} V_\alpha(t) < l$ for all $t \in M$. Let $N \subset M$ with $\operatorname{meas} N > 0$ be any subset. Set $p = 0$, $q = 0$ and

$$\overline{r}(t) \begin{cases} \in \{r \in \ker V_\alpha(t) \mid |r| = 1\} & \text{for } t \in N, \\ = 0 & \text{for } t \notin N, \end{cases} \tag{2.28}$$

where $\overline{r}(t)$ is chosen in such a way that $\overline{r}$ is a measurable function. We have

$$\dot{y} - Ay - Bv = 0, \tag{2.29}$$

$$\Xi_0 y(0) + \Xi_1 y(1) = 0, \tag{2.30}$$

$$\Upsilon y + \Theta v + T_\alpha \vartheta = \overline{r}. \tag{2.31}$$

By surjectivity, there exists a constant $k_s > 0$ and a solution $(\overline{y}, \overline{v}, \overline{\vartheta})$ of (2.29)–(2.31) such that

$$\|(\overline{y}, \overline{v}, \overline{\vartheta})\|_{Z^s} \le k_s \|\overline{r}\|_s.$$

Hence, in view of (2.28),

$$\|\overline{v}\|_s \le k_s (\operatorname{meas} N)^{1/s}. \tag{2.32}$$

On the other hand, in view of (2.28), equation (2.31) yields $|\Upsilon(t)\overline{y}(t)| = 1$ for $t \in N$, i.e., there exists $\overline{c} > 0$ such that

$$|\overline{y}(t)| \ge \overline{c} \quad \text{for } t \in N. \tag{2.33}$$

Note that, for any $v \in L^s(0,1;\mathbb{R}^m)$, there is a unique solution to (2.29)–(2.30). Indeed, using the definition (2.6), we obtain from (2.29)

$$y(t) = \Phi(t)y(0) + \int_0^t \Phi(t)\Phi(\tau)^{-1}B(\tau)v(\tau)\,d\tau. \tag{2.34}$$

By (2.30) and (2.34), we get

$$[\Xi_0 + \Xi_1\Phi(1)]\,y(0) = -\Xi_1 \int_0^1 \Phi(t)\Phi(\tau)^{-1}B(\tau)v(\tau)\,d\tau,$$

i.e., in view of (III), we have

$$y(0) = -\left[\Xi_0 + \Xi_1\Phi(1)\right]^{-1} \Xi_1 \int_0^1 \Phi(t)\Phi(\tau)^{-1}B(\tau)v(\tau)\,d\tau. \tag{2.35}$$

Substituting (2.35) into (2.34) we obtain

$$\begin{aligned} y(t) = &-\Phi(t)[\Xi_0 + \Xi_1\Phi(1)]^{-1}\Xi_1 \int_0^1 \Phi(t)\Phi(\tau)^{-1}B(\tau)v(\tau)\,d\tau \\ &+ \int_0^t \Phi(t)\Phi(\tau)^{-1}B(\tau)v(\tau)\,d\tau. \end{aligned} \tag{2.36}$$

Hence, there exists $c_s < \infty$ such that

$$\|y\|_\infty \le \|y\|_{1,s} \le c_s \|v\|_s. \tag{2.37}$$

Therefore, in view of (2.32), we have $\|\overline{y}\|_\infty \le c_s k_s (\operatorname{meas} N)^{1/s}$. Choosing N such that

$$\operatorname{meas} N \le \left(\frac{\overline{c}}{2c_s k_s}\right)^s$$

we contradict (2.33) and complete the proof of the lemma. ■

For a given $\alpha \ge 0$ introduce the sets

$$M_\alpha^i = \{t \in [0,1] \mid i \in I_\alpha(t)\} \quad \text{for } i \in I \tag{2.38}$$

and the spaces

$$L^s(M^i_\alpha;\mathbb{R}),\quad \widehat{L}^s_\alpha=\prod_{i\in I}L^s(M^i_\alpha;\mathbb{R}),\quad \widehat{Y}^s_\alpha=L^s(0,1;\mathbb{R}^n)\times\mathbb{R}^n\times\widehat{L}^s_\alpha. \tag{2.39}$$

Define the mapping

$$\widehat{\mathcal{C}}_\alpha:X^s\to\widehat{Y}^s_\alpha,\quad \widehat{\mathcal{C}}_\alpha\begin{pmatrix}y\\v\end{pmatrix}=\begin{pmatrix}\dot y-Ay-Bv\\ \varXi_0 y(0)+\varXi_1 y(1)\\ \widehat{\varUpsilon}_\alpha y+\widehat{\varTheta}_\alpha v\end{pmatrix}. \tag{2.40}$$

Using Lemma 2.1 and the same argument as in the proof of Lemma 2.3 we obtain:

COROLLARY 2.4. *Assumptions* (A1) *and* (A2) *imply that* $\widehat{\mathcal{C}}_\alpha:Z^s\to\widehat{Y}^s_\alpha$ *is surjective for any* $s\in[1,\infty]$. *If* $\widehat{\mathcal{C}}_\alpha:Z^s\to\widehat{Y}^s_\alpha$ *is surjective for any* $s\in[1,\infty)$ *then* (A1) *and* (A2) *hold.*

Bibliographical note. Constraint qualifications of the form (A1) with $\alpha=0$ can be found in [22]. In stability analysis, they were used in [1, 2]. For $\alpha>0$ these conditions were introduced in [34]. Conditions of the form (2.15) were used by W. W. Hager [19] in regularity analysis of solutions to optimal control problems subject to control and state constraints. For mixed constraints, they were exploited by V. Zeidan [54]. Controllability condition (A2) and its characterization by (2.17) were given in [34] and [39]. A very similar characterization is also given in [54]. Lemma 2.3 is based on a similar result in [39]. Lemma 2.1 was first proved by U. Felgenhauer (unpublished).

3. Optimality condition and coercivity

Let us introduce the following Lagrangian and Hamiltonian associated with $(\mathrm{O})_h$:

$$\begin{aligned}&\mathcal{L}:X^\infty\times(Y^\infty)^*\times\mathcal{G}\to\mathbb{R},\quad \mathcal{H}:\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^n\times\mathbb{R}^l\times G,\\ &\mathcal{L}(x,u,p,\varrho,\nu,h)=F(x,u,h)-(p,\dot x-f(x,u,h))\\ &\qquad\qquad+\langle\varrho,\xi(x(0),x(1),h(0)\rangle+(\nu,\theta(x,u,h)),\end{aligned} \tag{3.1}$$

$$\mathcal{H}(x,u,p,\nu,h)=\varphi(x,u,h)+\langle p,f(x,u,h)\rangle+\langle\nu,\theta(x,u,h)\rangle. \tag{3.2}$$

LEMMA 3.1. *If conditions* (A1) *and* (A2) *are satisfied, then there exists a unique Lagrange multiplier* $(p_0,\varrho_0,\nu_0)\in Y^{1,\infty}$ *such that the following first order optimality conditions in the Karush–Kuhn–Tucker* (*KKT*) *form are satisfied*:

$$\begin{cases}D_x\mathcal{L}(x_0,u_0,p_0,\varrho_0,\nu_0,h_0)=0,\\ D_u\mathcal{L}(x_0,u_0,p_0,\varrho_0,\nu_0,h_0)=0,\\ (\nu_0,\theta(x_0,u_0,h_0))=0,\quad \nu_0\in K_+,\end{cases} \tag{3.3}$$

where $K_+\subset(L^\infty(0,1;\mathbb{R}^l))^*$ *denotes the cone of linear functionals nonnegative on* $K=\{y\in L^\infty(0,1;\mathbb{R}^l)\mid y(t)\ge 0$ *for a.a.* $t\in[0,1]\}$.

Proof. First we show that there is a normal Lagrange multiplier in $(Y^\infty)^*$; then we will prove that it is more regular. The known *constraint regularity condition* (see, e.g., Theorem 1, Section 9.4 in [31]) implies that a normal Lagrange multiplier exists if there is a positive constant $a > 0$ and a pair $(y, v) \in X^\infty$ such that

$$
(3.4)\qquad \begin{cases} \dot{y}(t) - A(t)y(t) - B(t)v(t) = 0 & \text{for a.a. } t \in [0,1], \\ \Xi_0 y(0) + \Xi_1 y(1) = 0, \\ \theta_0^i(t) + \langle \Upsilon^i(t), y(t)\rangle + \langle \Theta^i(t), v(t)\rangle \le -a & \text{for all } i \in I \text{ and for a.a. } t \in [0,1]. \end{cases}
$$

Note that, by the Banach open mapping theorem (see, e.g., [47]), surjectivity of the mapping $\mathcal{C}_\alpha : Z^\infty \to Y^\infty$ implies that there exists a constant $k > 0$ such that for any $(p, q, r) \in Y^\infty$ there exists a solution (y, v, ϑ) of (2.21) such that

$$
(3.5)\qquad \|(y, v, \vartheta)\|_{Z^\infty} \le k\|(p, q, r)\|_{Y^\infty}.
$$

Choose $p = 0$, $q = 0$, $r(t) = -r\mathbf{1}$, where $r = \frac{1}{2}\alpha|k\|\theta_0\|_\infty - 1|^{-1}$ and $\mathbf{1} = (1, \ldots, 1)^* \in \mathbb{R}^l$. Let (y, v, ϑ) be a solution of (2.21) such that (3.5) holds. Since $\theta_0(t) \le 0$, in view of (2.9), we find from the last line in (2.21) that

$$
\theta_0^i(t) + \langle \Upsilon^i(t), y(t)\rangle + \langle \Theta^i(t), v(t)\rangle \le -r \quad \text{for } i \in I_\alpha(t).
$$

On the other hand, if $i \notin I_\alpha(t)$, then $\theta_0(t) \le -\alpha$. Hence, if we choose $r = \frac{1}{2}\alpha[k(\|\Upsilon\|_\infty + \|\Theta\|_\infty)]^{-1}$, we get

$$
\theta_0^i(t) + \langle \Upsilon^i(t), y(t)\rangle + \langle \Theta^i(t), v(t)\rangle \le -\tfrac{1}{2}\alpha \quad \text{for } i \notin I_\alpha(t).
$$

Thus (3.4) is satisfied with $a = \frac{1}{2}\alpha \min\{1, [k(\|\Upsilon\|_\infty + \|\Theta\|_\infty)]^{-1}\}$ and a Lagrange multiplier $(p_0, q_0, \nu_0) \in (Y^\infty)^*$ exists.

Let us rewrite the first two equations in (3.3) in an explicit form:

$$
(3.6)\qquad D_x\varphi(x_0, u_0, h_0) + \dot{p}_0 + A^* p_0 + D_x\Theta^*(x_0, u_0, h_0)\nu_0 = \dot{p}_0 + D_x\mathcal{H}(x_0, u_0, p_0, \nu_0, h_0) = 0,
$$

$$
(3.7)\qquad \begin{cases} p_0(0) + \Xi_0^* \varrho_0 + D_{x(0)}\psi(x_0(0), x_0(1), h_0(0)) = 0, \\ -p_0(1) + \Xi_1^* \varrho_0 + D_{x(1)}\psi(x_0(0), x_0(1), h_0(0)) = 0, \end{cases}
$$

$$
(3.8)\qquad D_u\varphi(x_0, u_0, h_0) + B^* p_0 + D_u\Theta^*(x_0, u_0, h_0)\nu_0 = D_u\mathcal{H}(x_0, u_0, p_0, \nu_0, h_0) = 0.
$$

In view of definition (2.9), the last equation in (3.3) implies

$$
(3.9)\qquad T_\alpha^* \nu_0 = 0.
$$

Using definition (2.20), we can rewrite (3.6)–(3.9) in the following compact form:

$$
(3.10)\qquad \mathcal{C}_\alpha^* \begin{bmatrix} p_0 \\ \varrho_0 \\ \nu_0 \end{bmatrix} = \kappa_0,
$$

where $\mathcal{C}_\alpha^* : (Y^s)^* \to (Z^s)^*,\ s \in [2, \infty]$, is the operator adjoint to $\mathcal{C}_\alpha$, and

$$
(3.11)\qquad \kappa_0 = \begin{bmatrix} -D_x\varphi_0 \\ -p_0(0) - D_{x(0)}\psi_0 \\ p_0(1) - D_{x(1)}\psi_0 \\ -D_u\varphi_0 \\ 0 \end{bmatrix}.
$$

Clearly, by (I), $\kappa_0 \in (Z^2)^*$. Since Z^2 is a Hilbert space, there exists a canonical isomorphism $\mathcal{J} : (Z^2)^* \to Z^2$, and from (3.10) we get

$$\mathcal{J}\mathcal{C}_\alpha^* \begin{bmatrix} p_0 \\ \varrho_0 \\ \nu_0 \end{bmatrix} = \mathcal{J}\kappa_0, \tag{3.12}$$

where $\mathcal{J}\kappa_0 \in Z^2$. In view of Lemma 2.3, $\mathcal{J}\mathcal{C}_\alpha^* : (Y^2)^* = Y^2 \to Z^2$ is injective and $\mathcal{C}_\alpha\mathcal{J}\mathcal{C}_\alpha^* : Y^2 \to Y^2$ is invertible. So, (3.12) yields

$$\begin{bmatrix} p_0 \\ \varrho_0 \\ \nu_0 \end{bmatrix} = (\mathcal{C}_\alpha\mathcal{J}\mathcal{C}_\alpha^*)^{-1}\,\mathcal{C}_\alpha\mathcal{J}\kappa_0, \tag{3.13}$$

which shows that (p_0, ϱ_0, ν_0) is unique and belongs to Y^2. In particular $\nu_0 \in L^2(0,1;\mathbb{R}^n)$.

We show that actually $(p_0, \varrho_0, \nu_0) \in Y^{1,\infty}$. To this end, note that, since ν_0 is a function, (3.9) implies

$$T_\alpha(t)^*\nu_0(t) = 0 \quad \text{for a.a. } t \in [0,1].$$

Combining this equation with (3.8) and using definition (2.9), we obtain

$$V_\alpha(t)^*\nu_0(t) = \begin{bmatrix} -(D_u\varphi_0(t) + B^*(t)p_0(t)) \\ 0 \end{bmatrix},$$

and, by (A1),

$$|\nu_0(t)| \le \beta^{-1} \left| V_\alpha(t) \begin{bmatrix} (D_u\varphi_0(t) + B^*(t)p_0(t)) \\ 0 \end{bmatrix} \right|.$$

Hence, in view of (I), $\nu_0 \in L^\infty(0,1;\mathbb{R}^l)$ and by (3.6), $p_0 \in W^{1,\infty}(0,1;\mathbb{R}^n)$. ■

Note that since $\nu_0 \in L^\infty(0,1;\mathbb{R}^l)$, the last condition in (3.3) yields the following pointwise complementarity:

$$\langle \nu_0(t), \theta(x_0(t), u_0(t), h_0(t))\rangle = 0, \quad \nu_0(t) \ge 0 \quad \text{for a.a. } t \in [0,1]. \tag{3.14}$$

We still need a coercivity condition. To this end, for $\alpha \ge 0$ we introduce the sets

$$J_\alpha(t) = \{i \in I_0(t) \mid \nu_0^i(t) > \alpha\}, \tag{3.15}$$

and, as in (2.38), (2.39), define

$$\begin{cases} N_\alpha^i = \{t \in [0,1] \mid i \in J_\alpha(t)\} \quad \text{for } i \in I, \\ L^s(N_\alpha^i;\mathbb{R}), \quad \overline{L}_\alpha^s = \prod_{i\in I} L^s(N_\alpha^i;\mathbb{R}), \quad \overline{Y}_\alpha^s = L^s(0,1;\mathbb{R}^n) \times \mathbb{R}^n \times \overline{L}_\alpha^s. \end{cases} \tag{3.16}$$

We assume

(A3) (*Coercivity*) There exist constants $\alpha, \gamma > 0$ such that

$$\begin{aligned} ((y,v), D^2\mathcal{L}_0(y,v)) := \int_0^1 \begin{bmatrix} y(t) \\ v(t) \end{bmatrix}^* \begin{bmatrix} Q_{11}(t) & Q_{12}(t) \\ Q_{21}(t) & Q_{22}(t) \end{bmatrix} \begin{bmatrix} y(t) \\ v(t) \end{bmatrix} dt \\ + \begin{bmatrix} y(0) \\ y(1) \end{bmatrix}^* \begin{bmatrix} \mathcal{R}_{00} & \mathcal{R}_{01} \\ \mathcal{R}_{10} & \mathcal{R}_{11} \end{bmatrix} \begin{bmatrix} y(0) \\ y(1) \end{bmatrix} \\ \ge \gamma\|v\|_2^2 \end{aligned} \tag{3.17}$$

for all $(y,v) \in X^2$ such that

$$(y,v) \in \ker \overline{\mathcal{C}}_\alpha, \tag{3.18}$$

where

$$(3.19)\qquad \begin{cases} Q_{11} = D^2_{xx}\mathcal{H}_0, \quad Q_{12} = D^2_{xu}\mathcal{H}_0, \quad Q_{21} = D^2_{ux}\mathcal{H}_0, \quad Q_{22} = D^2_{uu}\mathcal{H}_0, \\ \mathcal{R}_{ij} = D^2_{x(i)x(j)}(\xi(x_0(0), x_0(1), h_0(0))^*\varrho_0 + \psi(x_0(0), x_0(1), h_0(0))), \\ i, j = 0, 1, \end{cases}$$

$$(3.20)\qquad \overline{\mathcal{C}}_\alpha \begin{bmatrix} y \\ v \end{bmatrix} = \begin{bmatrix} \dot{y} - Ay - Bv \\ \varXi_0 y(0) + \varXi_1 y(1) \\ \overline{\varUpsilon}_\alpha y + \overline{\varTheta}_\alpha v \end{bmatrix},$$

$$(3.21)\qquad \overline{\varTheta}_\alpha(t) = [D_u\theta_0^i(t)]_{i\in J_\alpha(t)}, \quad \overline{\varUpsilon}_\alpha(t) = [D_x\theta_0^i(t)]_{i\in J_\alpha(t)}.$$

Note that, since (y, v) is a solution to (2.29) and (2.30), the estimate (2.37) holds and (3.17) implies

$$(3.22)\qquad ((y,v), D^2\mathcal{L}_0(y,v)) \geq \gamma'(\|y\|^2_{1,2} + \|v\|^2_2) \quad \text{for all } (y,v) \in \ker \overline{\mathcal{C}}_\alpha.$$

where $\gamma' = \gamma(1 + c_2^2)^{-1}$.

Define the following continuous mappings:

$$(3.23)\qquad \begin{aligned} &\mathcal{S} : L^s(0,1;\mathbb{R}^n) \to W^{1,s}(0,1;\mathbb{R}^n) \times \mathbb{R}^n \times \mathbb{R}^n, \\ &\mathcal{S}_0 : L^s(0,1;\mathbb{R}^n) \to W^{1,s}(0,1;\mathbb{R}^n), \\ &\mathcal{S}k = (y, y(0), y(1)), \quad \mathcal{S}_0 k = y, \end{aligned}$$

where $s \in [1,\infty]$ and y is the solution to

$$\dot{y}(t) - A(t)y(t) - k(t) = 0, \quad \varXi_0 y(0) + \varXi_1 y(1) = 0,$$

i.e., as in (2.36), we have

$$y(t) = -\varPhi(t)[\varXi_0 + \varXi_1\varPhi(1)]^{-1}\varXi_1 \int_0^1 \varPhi(t)\varPhi(\tau)^{-1}k(\tau)\,d\tau + \int_0^t \varPhi(t)\varPhi(\tau)^{-1}k(\tau)\,d\tau.$$

Using the mappings (3.23), we can eliminate y from (3.17) and express Coercivity (A3) as

$$(3.24)\qquad (u, (\mathcal{M} + Q_{22})u) \geq \gamma\|u\|_2^2 \quad \text{for all } u \in \mathcal{U}^2_\alpha,$$

where

$$(3.25)\qquad \mathcal{M} = B^*\mathcal{S}^* \begin{pmatrix} Q_{11} & 0 & 0 \\ 0 & \mathcal{R}_{00} & \mathcal{R}_{01} \\ 0 & \mathcal{R}_{10} & \mathcal{R}_{11} \end{pmatrix} \mathcal{S}B + B^*\mathcal{S}_0^* Q_{12} + Q_{21}\mathcal{S}_0 B,$$

and

$$(3.26)\qquad \mathcal{U}^s_\alpha = \{v \in L^s(0,1;\mathbb{R}^m) \mid \overline{\varUpsilon}_\alpha(t)(\mathcal{S}_0 v)(t) + \overline{\varTheta}_\alpha(t)v(t) = 0\}$$

is a closed subspace of $L^s(0,1;\mathbb{R}^m)$. If we denote by

$$(3.27)\qquad \varGamma_\alpha : L^2(0,1;\mathbb{R}^m) \to \mathcal{U}^2_\alpha \quad \text{the orthogonal projection onto } \mathcal{U}^2_\alpha,$$

then (A3) can be expressed in the form

$$(3.28)\qquad \begin{cases} (v, \mathcal{Q}_\alpha v) \geq \gamma\|\varGamma_\alpha v\|_2^2 & \text{for all } v \in L^2(0,1;\mathbb{R}^m) \quad \text{or} \\ (v, \mathcal{Q}_\alpha v) \geq \gamma\|v\|_2^2 & \text{for all } v \in \mathcal{U}^2_\alpha, \end{cases}$$

where

$$(3.29)\qquad \mathcal{Q}_\alpha = \varGamma_\alpha(\mathcal{M} + Q_{22})\varGamma_\alpha.$$

REMARK 3.2. The coercivity condition (3.17) holds in the *weaker* norm of the space X^2, rather than in X^∞, in which problem $(\mathrm{O})_h$ is well defined and the Lagrangian is twice differentiable. In X^2 this differentiability property is not satisfied. This phenomenon is called *two-norm discrepancy* and it is typical of nonlinear optimal control problems [23, 39].

LEMMA 3.3. *If* (A1) *and* (A2) *hold*, *then* (A3) *implies the following Legendre–Clebsch condition*:

$$\langle v, Q_{22}(t)v\rangle \geq \gamma|u|^2 \quad \text{for all } v \in \{\mathbb{R}^m \mid \langle \Theta^i(t), v\rangle = 0 \text{ for all } i \in J_\alpha(t)\} \text{ and for a.a. } t \in [0,1]. \tag{3.30}$$

Proof. Suppose that (3.30) is violated, i.e., there exist a set $M \subset [0,1]$ with $\operatorname{meas} M > 0$, a constant $\varepsilon > 0$ and a vector $\check{v}(t) \in \mathbb{R}^m$, $|\check{v}(t)| = 1$, satisfying the conditions $\langle \Theta^i(t), \check{v}(t)\rangle = 0$ for $i \in J_\alpha(t)$, such that

$$\langle \check{v}(t), Q_{22}(t)\check{v}(t)\rangle \leq \gamma - \varepsilon \quad \text{for all } t \in M. \tag{3.31}$$

We can choose $\check{v}(t)$ in such a way that the function $\check{v}(\cdot)$ is measurable on M. For an arbitrary subset $N \subset M$ define the function

$$\widehat{v}(t) = \begin{cases} \check{v}(t) & \text{for } t \in N, \\ 0 & \text{for } t \notin N. \end{cases} \tag{3.32}$$

Let $\widehat{y}$ be the corresponding solution to (2.29), (2.30). By (2.36) and (3.32) there exists a constant c such that

$$\|\widehat{y}\|_\infty \leq c \operatorname{meas} N. \tag{3.33}$$

Let us put $\overline{r} = \Upsilon\widehat{y}$ in (2.31). By Lemma 2.3 and by the Banach open mapping theorem, there exists a solution $(\overline{y}, \overline{v}, \overline{\vartheta})$ of (2.29)–(2.31) and a constant $k > 0$ such that

$$\|(\overline{y}, \overline{v}, \overline{\vartheta})\|_{Z^\infty} \leq k\|(0,0,\overline{r})\|_{Y^\infty}.$$

Hence, in view of (3.33), there exists $c_1 > 0$ such that

$$\|(\overline{y}, \overline{v}, \overline{\vartheta})\|_{Z^\infty} \leq c_1 \operatorname{meas} N. \tag{3.34}$$

Clearly, $(\widetilde{y}, \widetilde{v}) := (\widehat{y} - \overline{y}, \widehat{v} - \overline{v})$ satisfies (2.29)–(2.30). Note that, for $i \in J_\alpha(t) \subset I_0(t)$ we have $T^i_\alpha(t) = 0$. Hence, it follows from the construction that

$$\langle \Upsilon^i(t), \widetilde{y}(t)\rangle + \langle \Theta^i(t), \widetilde{v}(t)\rangle = 0 \quad \text{for } i \in J_\alpha(t),$$

i.e., $(\widetilde{y}, \widetilde{v})$ satisfies (3.18). We are going to show that (3.17) is violated by $(\widetilde{y}, \widetilde{v})$, provided that $\operatorname{meas} N$ is sufficiently small. This will contradict (3.31) and complete the proof of the lemma. Indeed, by (3.32) and (3.34) we have

$$\|\widetilde{v}\|_2^2 = \|\widehat{v} - \overline{v}\|_2^2 \geq \operatorname{meas} N(1 - c_1 \operatorname{meas} N)^2 \geq \operatorname{meas} N - 2c_1(\operatorname{meas} N)^2. \tag{3.35}$$

On the other hand, (3.31)–(3.34) imply that

$$((\widetilde{y}, \widetilde{v}), D^2\mathcal{L}_0(\widetilde{y}, \widetilde{v})) \leq (\gamma - \varepsilon) \operatorname{meas} N + c_2(\operatorname{meas} N)^2, \tag{3.36}$$

where $c_2 > 0$ is independent of the choice of $N \subset M$. It follows from (3.35) and (3.36) that, if we choose N such that

$$\operatorname{meas} N \leq \frac{\varepsilon}{2}[c_2 + 2c_1(\gamma - \varepsilon/2)]^{-1}$$

then

$$((\widetilde{y},\widetilde{v}), D^2\mathcal{L}_0(\widetilde{y},\widetilde{v})) \le (\gamma-\varepsilon/2)\|\widetilde{v}\|_2^2 \le (\gamma-\varepsilon/2)(\|\widetilde{y}\|_{1,2}^2 + (\|\widetilde{v}\|_2^2),$$

which violates (3.17) and completes the proof of the lemma. ∎

Bibliographical note. Necessary and sufficient optimality conditions for problems with mixed control-state constraints were derived and thoroughly discussed in [54]. Sufficient optimality conditions for such problems were also obtained in [41]. Coercivity condition (A3) with $\alpha > 0$ was used in [15], where also Lemma 3.3 was proved. The important phenomenon of *two-norm discrepancy* in optimal control was first described in [23] and [40].

4. Application of abstract theorems

Combining (2.2)–(2.4) with (3.6)–(3.8) and (3.14), evaluated at h rather than at h_0, we obtain the following optimality system for $(\mathrm{O})_h$:

$$(4.1)\qquad \begin{cases} \dot{x}(t) - f(x(t),u(t),h(t)) = 0, \\ \xi(x(0),x(1),h(0)) = 0, \\ \dot{p}(t) + D_x\mathcal{H}(x(t),u(t),p(t),\nu(t),h(t)) = 0, \\ p(0) + D_{x(0)}[\xi(x(0),x(1),h(0))^*\varrho + \psi(x(0),x(1),h(0))] = 0, \\ -p(1) + D_{x(1)}[\xi(x(0),x(1),h(0))^*\varrho + \psi(x(0),x(1),h(0))] = 0, \\ D_u\mathcal{H}(x(t),u(t),p(t),\nu(t),h(t)) = 0, \end{cases}$$

$$(4.2)\qquad \begin{cases} \langle \nu(t), \theta(x(t),u(t),h(t))\rangle = 0, \\ \theta(x(t),u(t),h(t)) \le 0, \quad \nu(t) \ge 0. \end{cases}$$

It will be convenient to express (4.2) in the form of an inclusion. To this end, we define the normal cone to the positive orthant $\mathbb{R}^l_+ \subset \mathbb{R}^l$ by

$$(4.3)\qquad N_{\mathbb{R}^l_+}(\nu) = \begin{cases} \{y \in \mathbb{R}^l \mid \langle y, \mu - \nu\rangle \le 0 \ \forall \mu \in \mathbb{R}^l_+\} & \text{if } \nu \in \mathbb{R}^l_+, \\ \emptyset & \text{if } \nu \notin \mathbb{R}^l_+. \end{cases}$$

In terms of (4.3), we can express (4.2) in the following equivalent form:

$$(4.4)\qquad \theta(x(t),u(t),h(t)) \in N_{\mathbb{R}^l_+}(\nu(t)) \quad \text{for a.a. } t \in [0,1].$$

Using the definitions (2.39) and (3.16), we introduce the spaces

$$\begin{aligned} W^s &= X^s \times Y^{1,s}, \\ U^s &= L^s(0,1;\mathbb{R}^n) \times \mathbb{R}^n \times L^s(0,1;\mathbb{R}^l) \times L^s(0,1;\mathbb{R}^n) \times \mathbb{R}^n \times \mathbb{R}^n \times L^s(0,1;\mathbb{R}^m), \\ \widehat{U}^s_\alpha &= L^s(0,1;\mathbb{R}^n) \times \mathbb{R}^n \times \widehat{L}^s_\alpha \times L^s(0,1;\mathbb{R}^n) \times \mathbb{R}^n \times \mathbb{R}^n \times L^s(0,1;\mathbb{R}^m), \\ \overline{U}^s_\alpha &= L^s(0,1;\mathbb{R}^n) \times \mathbb{R}^n \times \overline{L}^s_\alpha \times L^s(0,1;\mathbb{R}^n) \times \mathbb{R}^n \times \mathbb{R}^n \times L^s(0,1;\mathbb{R}^m), \end{aligned}$$

for $s \in [1,\infty]$. The optimality system (4.1), (4.4) can be rewritten in the form of the following *generalized equation*:

$$(4.5)\qquad 0 \in \mathcal{F}(\zeta,h) + \mathcal{T}(\zeta),$$

where $\zeta = (x, u, p, \varrho, \nu) \in W^\infty$, $\mathcal{F} : W^\infty \times \mathcal{G} \to U^\infty$ is the function

$$\mathcal{F}(\zeta, h) = \begin{pmatrix} \dot{x} - f(x, u, h) \\ \xi(x(0), x(1), h(0)) \\ \theta(x, u, h) \\ \dot{p} + D_x\mathcal{H}(x, u, p, \nu, h) \\ p(0) + D_{x(0)}[\xi(x(0), x(1), h(0))^*\varrho + \psi(x(0), x(1), h(0))] \\ -p(1) + D_{x(1)}[\xi(x(0), x(1), h(0))^*\varrho + \psi(x(0), x(1), h(0))] \\ D_u\mathcal{H}(x, u, p, \nu, h) \end{pmatrix}, \tag{4.6}$$

and $\mathcal{T} : W^\infty \to 2^{U^\infty}$ is the multivalued mapping with closed graph given by

$$\mathcal{T}(\zeta) = \begin{pmatrix} 0 \\ 0 \\ -\mathcal{N}(\nu) \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{4.7}$$

with $(\mathcal{N}(\nu))(t) := N_{\mathbb{R}^l_+}(\nu(t))$.

We are going to investigate conditions under which there are neighborhoods $\mathcal{G}_0 \subset \mathcal{G}$ and $\mathcal{B}_0 \subset W^\infty$ of h_0 and ζ_0, respectively, such that for each $h \in \mathcal{G}_0$ there exists a unique solution to (4.1), (4.2) (or equivalently to (4.1), (4.4)) in $\mathcal{B}_0$ and it is a Lipschitz continuous and differentiable (in some sense) function of h. To this end, we will recall some results concerning abstract generalized equations. Along with (4.5), let us introduce the following generalized equation obtained by linearization and perturbation of (4.5) at the reference point ζ_0:

$$\delta \in \mathcal{F}(\zeta_0, h_0) + D_\zeta\mathcal{F}(\zeta_0, h_0)(\eta - \zeta_0) + \mathcal{T}(\eta), \tag{4.8}$$

where $\delta \in U^\infty$ is a perturbation. Clearly, for $\delta = 0$,

$$\eta_0 = \zeta_0 \tag{4.9}$$

is a solution to (4.8).

In the analysis of Lipschitz stability of solutions to (4.5), a crucial role is played by Robinson's implicit function theorem for generalized equations (see Theorem 2.1 and Corollary 2.2 in [44]), which for our purpose can be formulated as follows.

THEOREM 4.1. *Suppose that $\mathcal{F}(\cdot,\cdot)$ is Fréchet differentiable in a neighborhood of (ζ_0, h_0) and $D_\zeta\mathcal{F}(\cdot,\cdot)$ is Lipschitz continuous in both variables. If there exists a constant $\widehat{l}$ such that*

(i) *for any $\varepsilon > 0$ there exist $\varsigma_1, \varsigma_2 > 0$ such that for each $\delta \in \mathcal{B}^{U^\infty}_{\varsigma_1}(0)$ there is a unique solution η_δ in $\mathcal{B}^{W^\infty}_{\varsigma_2}(\zeta_0)$ to the linearized general equation* (4.8), *and it is Lipschitz continuous in δ with modulus $\widehat{l} + \varepsilon$,*

then

(ii) *for any $\varepsilon > 0$ there exist $\sigma_1, \sigma_2 > 0$ such that for each $h \in \mathcal{B}^H_{\sigma_1}(h_0)$ there is a unique solution ζ_h in $\mathcal{B}^{W^\infty}_{\sigma_2}(\zeta_0)$ to the nonlinear generalized equation* (4.5), *and it is Lipschitz continuous in h with modulus $\widehat{l} + \varepsilon$.*

Note that in Robinson's theorem condition (i) plays the role analogous to regularity of the Jacobian in the classical implicit function theorem.

In sensitivity analysis we will use several notions of differentiability. Let us recall some of them.

DEFINITION 4.2. Let H and X be Banach spaces. A function $\phi : H \to X$ is called *directionally differentiable* at h_0 if for every $g \in H$ there exists $D_h\phi(h_0; g) \in X$ with the property that for every $\varepsilon > 0$ there exists $\tau > 0$ such that

$$\|\phi(h_0 + tg) - \phi(h_0) - tD_h\phi(h_0; g)\|_X \le t\varepsilon \quad \text{for every } t \in (0, \tau]. \tag{4.10}$$

If (4.10) holds for every $t \in [-\tau, \tau]$, then ϕ is called *Gateaux differentiable* at h_0.

If there exists a positively homogeneous mapping $D_h\phi(h_0) : H \to X$ with the property that for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\|\phi(h) - \phi(h_0) - (D_h\phi(h_0), h - h_0)\|_X \le \varepsilon\|h - h_0\|_H \quad \text{for } \|h - h_0\|_H < \delta, \tag{4.11}$$

then ϕ is called *Bouligand differentiable* (or *B-differentiable*) at h_0 and $D_h\phi(h_0)$ is called the *B-derivative*. If $D_h\phi(h_0) : H \to X$ is linear and bounded, the function ϕ is *Fréchet differentiable*. A Fréchet differentiable function is *strictly Fréchet differentiable* at h_0 if for every ε there exists $\delta > 0$ such that

$$\|\phi(h_1) - \phi(h_2) - (D_h\phi(h_0), h_1 - h_2)\|_H \le \varepsilon\|h_1 - h_2\|_H \quad \text{whenever } h_1, h_2 \in \mathcal{B}^H_\delta(h_0). \tag{4.12}$$

We will need the Lyusternik–Graves theorem, which is formulated below as in [11] and [24].

THEOREM 4.3. *Let H and X be Banach spaces and let $\phi : H \to X$ be a function strictly Fréchet differentiable at h_0 and $\phi(h_0) = x_0$. Let the derivative $D_h\phi(h_0) : H \to X$ be surjective. Then there exist constants $\sigma, \kappa > 0$ such that for every $x \in \mathcal{B}^X_\sigma(x_0)$, the equation $\phi(h) = x$ has a solution $h \in \mathcal{B}^H_{\kappa\sigma}(h_0)$. Moreover, there exist $\varrho, k > 0$ such that, for any $h \in \mathcal{B}^H_\varrho(h_0)$, there is $\widetilde{h} \in H$ with the property that*

$$\phi(\widetilde{h}) = \phi(h_0) \quad \textit{and} \quad \|h - \widetilde{h}\|_H \le k\|\phi(h) - \phi(h_0)\|_X.$$

In our sensitivity analysis for solutions to (4.5), we will use the differentiability part of Dontchev's implicit function theorem (see Theorem 2.4, Remark 2.6 and Corollary 2.10 in [9]), which can be formulated as follows.

THEOREM 4.4. *Suppose that the assumptions of Theorem 4.1 hold. In addition, assume that the solution η_δ to the linearized equation (4.8) is a directionally (respectively, Gateaux, Bouligand, Fréchet) differentiable function of δ in a neighborhood of 0, with the differential at 0 in a direction $\delta \in U^\infty$ denoted by $(D_\delta\eta_0; \delta)$. Then the solution ζ_h to (4.5) is a directionally (respectively, Gateaux, Bouligand, Fréchet) differentiable function of h in a neighborhood of h_0 and the differential at h_0 is given by*

$$(D_h\zeta_0; g) = (D_\delta\eta_0; -D_h\mathcal{F}(\zeta_0, h_0)g) \quad \textit{for all } g \in H. \tag{4.13}$$

REMARK 4.5. In Theorem 4.1, Lipschitz continuity of η and ζ is understood in the sense of the same norm of the space W^∞ in which $\mathcal{F}(\cdot, h)$ is differentiable. On the other hand, Theorem 4.4 remains true if the differentiability is satisfied in a norm in the image space *weaker* than that in which Lipschitz continuity in Theorem 4.1 holds (see Remark 2.11 in [9]); e.g., in W^s $(s < \infty)$, rather than in W^∞. This property will be used in Section 6.

Theorems 4.1 and 4.4 allow one to deduce existence, local uniqueness, Lipschitz continuity and differentiability of solutions to (4.5) from the same properties of the solutions to the linearized generalized equation (4.8). Usually this last equation is much easier to analyze than the original one.

In order to apply these theorems to our problem $(\mathrm{O})_h$, we have to find the form of the linearized equation (4.8) for $\mathcal{F}$ and $\mathcal{T}$ given by (4.6) and (4.7), respectively. Let us set $\eta = (y, v, q, \varrho, \mu) \in W^\infty$, $\delta = (d^1, d^2, d^3, e^1, e^2, e^3, e^4) \in U^\infty$. By a direct computation we obtain

$$\begin{cases} \dot{y} - Ay - Bv + a^1 - d^1 = 0, \\ \varXi_0 y(0) + \varXi_1 y(1) + a^2 - d^2 = 0, \end{cases} \tag{4.14}$$

$$\varUpsilon y + \varTheta v + a^3 - d^3 \in \mathcal{N}(\mu), \tag{4.15}$$

$$\begin{cases} \dot{q} + A^* q + Q_{11} y + Q_{12} v + \varUpsilon^* \mu + b^1 - e^1 = 0, \\ q(0) + \mathcal{R}_{00} y(0) + \mathcal{R}_{01} y(1) + \varXi_0^* \varrho + b^2 - e^2 = 0, \\ -q(1) + \mathcal{R}_{10} y(0) + \mathcal{R}_{11} y(1) + \varXi_1^* \varrho + b^3 - e^3 = 0, \end{cases} \tag{4.16}$$

$$Q_{12} y + Q_{22} v + B^* q + \varTheta^* \mu + b^4 - e^4 = 0, \tag{4.17}$$

where

$$\begin{cases} a^1 = -(\dot{x}_0 - Ax_0 - Bu_0), \\ a^2 = -(\varXi_0 x_0(0) + \varXi_1 x_0(1)), \\ a^3 = \theta_0 - (\varUpsilon x_0 + \varTheta u_0), \\ b^1 = -(\dot{p}_0 + A^* p_0 + Q_{11} x_0 + Q_{12} u_0 + \varUpsilon^* \nu_0), \\ b^2 = -(p_0(0) + \mathcal{R}_{00} x_0(0) + \mathcal{R}_{01} x_0(1) + \varXi_0^* \varrho_0), \\ b^3 = -(-p_0(1) + \mathcal{R}_{10} x_0(0) + \mathcal{R}_{11} x_0(1) + \varXi_1^* \varrho_0), \\ b^4 = -(Q_{12} x_0 + Q_{22} u_0 + B^* p_0 + \varTheta^* \nu_0). \end{cases} \tag{4.18}$$

Clearly, as in (4.9),

$$\eta_0 := (y_0, v_0, q_0, \varrho_0, \mu_0) = (x_0, u_0, p_0, \varrho_0, \nu_0) \tag{4.19}$$

is a solution of (4.14)–(4.17) for $\delta = 0$.

An inspection shows that (4.14)–(4.17) can be interpreted as the stationarity condition (optimality system) for the following linear-quadratic accessory problem depending on the parameter δ:

$(\mathrm{LO})_\delta$ Find $(y_\delta, v_\delta) \in X^\infty$ such that

$$\mathcal{I}(y_\delta, v_\delta, \delta) = \min \mathcal{I}(y, v, \delta)$$

subject to

$$\dot{y}(t) - A(t)y(t)) - B(t)v(t) + a^1(t) - d^1(t) = 0, \tag{4.20}$$

$$\varXi_0 y(0) + \varXi_1 y(1) + a^2 - d^2 = 0, \tag{4.21}$$

$$\varUpsilon(t)y(t) + \varTheta(t)v(t) + a^3(t) - d^3(t) \le 0, \tag{4.22}$$

where

$$\mathcal{I}(y,v,\delta) = \tfrac{1}{2}((y,v), D^2\mathcal{L}_0(y,v)) + \int_0^1 [\langle b^1(t)-e^1(t), y(t)\rangle + \langle b^4(t)-e^4(t), v(t)\rangle]\,dt + \langle b^2-e^2, y(0)\rangle + \langle b^3-e^3, y(1)\rangle. \tag{4.23}$$

In view of Theorems 4.1 and 4.4, stability and sensitivity of the stationary points of the accessory problem $(\mathrm{LO})_\delta$, with respect to perturbations δ, are crucial to get the same properties for the original problem $(\mathrm{O})_h$. In the next two sections, we will concentrate on the analysis of the accessory problem.

Bibliographical note. Robinson's Theorem 4.1 was proved in [44]. Another implicit function theorem of Robinson, based on the concept of *strong approximation*, is given in [46]. An important generalization of this theorem due to A. L. Dontchev can be found in [9]. The concept of Bouligand differentiability, sometimes called *directional Fréchet* differentiability, was first used in sensitivity analysis by A. Shapiro [48] and S. M. Robinson [45]. Theorem 4.4 is due to A. L. Dontchev [9]. The abstract Theorem 4.1 was used for the first time in the stability analysis of optimal control problems in [25]. Numerous further applications can be found e.g., in [1, 2, 8, 13, 15, 14, 32, 33, 34, 35].

5. Stability analysis for the accessory problem

In this section we show that, for δ sufficiently small, $(\mathrm{LO})_\delta$ has a locally unique stationary point, which is a Lipschitz continuous function of δ. It will be convenient to introduce slack variables $\pi \in L^2(0,1;\mathbb{R}^l)$ and to define the following problem $(\widetilde{\mathrm{LO}})_\delta$, which is a modification of $(\mathrm{LO})_\delta$:

$(\widetilde{\mathrm{LO}})_\delta$ Find $(\widetilde{y}_\delta, \widetilde{v}_\delta, \widetilde{\pi}_\delta) \in Z^2$ such that

$$\widetilde{\mathcal{I}}(\widetilde{y}_\delta, \widetilde{v}_\delta, \widetilde{\pi}_\delta, \delta) = \min\left\{\widetilde{\mathcal{I}}(y,v,\pi,\delta) := \mathcal{I}(y,v,\delta) + \tfrac{1}{2}(\pi,\pi)\right\} \tag{5.1}$$

subject to

$$\dot{y}(t) - A(t)y(t)) - B(t)v(t) + a^1(t) - d^1(t) = 0, \tag{5.2}$$

$$\Xi_0 y(0) + \Xi_1 y(1) + a^2 - d^2 = 0, \tag{5.3}$$

$$\langle \Upsilon^i(t), y(t)\rangle + \langle \Theta^i(t), v(t)\rangle + \langle T^i_\alpha(t), \pi(t)\rangle + (a^3)^i(t) - (d^3)^i(t) \begin{cases} = 0 & \text{if } i \in J_\alpha(t), \\ \le 0 & \text{if } i \in I_\alpha(t) \setminus J_\alpha(t), \\ \text{free} & \text{if } i \notin I_\alpha(t). \end{cases} \tag{5.4}$$

Note that problem $(\widetilde{\mathrm{LO}_\delta})$ is defined on the Hilbert space Z^2, rather than on Z^∞. The constraints are modified in such a way that their linear part is given by $\mathcal{C}_\alpha$, so by Lemma 2.3 it is a surjective mapping from Z^2 into Z^2. Finally by (A3), the modified cost functional $\widetilde{\mathcal{I}}$ is coercive on the linear hull of the feasible set. Thanks to the above properties, the stability analysis for $(\widetilde{\mathrm{LO}_\delta})$ becomes simple. Later on, we will show that, for δ sufficiently small, the solutions $(\widetilde{y}_\delta, \widetilde{v}_\delta)$ to $(\widetilde{\mathrm{LO}}_\delta)$ coincide with the solutions (y_δ, v_δ) of (LO_δ) and thus we will obtain the required properties of the latter.

The optimality system for $(\widetilde{\mathrm{LO}}_\delta)$ is given by the following modification of (4.14)–(4.17):

$$\begin{aligned} &\dot{y}(t) - A(t)y(t) - B(t)v(t) + a^1(t) - d^1(t) = 0, \\ &\Xi_0 y(0) + \Xi_1 y(1) + a^2 - d^2 = 0, \end{aligned} \tag{5.5}$$

$$\langle \Upsilon^i(t), y(t)\rangle + \langle \Theta^i(t), v(t)\rangle + \langle T^i_\alpha(t), \pi(t)\rangle + (a^3)^i(t) - (d^3)^i(t) \begin{cases} = 0 & \text{if } i \in J_\alpha(t), \\ \le 0 & \text{if } i \in I_\alpha(t) \setminus J_\alpha(t), \\ \text{free} & \text{if } i \notin I_\alpha(t), \end{cases} \tag{5.6}$$

$$\begin{aligned} &\langle \mu(t), \Upsilon(t)y(t) + \Theta(t)v(t) + T_\alpha(t)\pi(t) + a^3(t) - d^3(t)\rangle = 0, \\ &\mu^i(t) \begin{cases} \ge 0 & \text{if } i \in I_\alpha(t) \setminus J_\alpha(t), \\ = 0 & \text{if } i \notin I_\alpha(t), \end{cases} \end{aligned} \tag{5.7}$$

$$\begin{aligned} &\dot{q}(t) + A^*(t)q + Q_{11}(t)y(t) + Q_{12}(t)v(t) + \Upsilon^*(t)\mu(t) + b^1(t) - e^1(t) = 0, \\ &q(0) + \mathcal{R}_{00}y(0) + \mathcal{R}_{01}y(1) + \Xi_0^*\varrho + b^2 - e^2 = 0, \\ &-q(1) + \mathcal{R}_{10}y(0) + \mathcal{R}_{11}y(1) + \Xi_1^*\varrho + b^3 - e^3 = 0, \end{aligned} \tag{5.8}$$

$$Q_{21}(t)y(t) + Q_{22}(t)v(t) + B^*(t)q(t) + \Theta^*(t)\mu(t) + b^4(t) - e^4(t) = 0, \tag{5.9}$$

$$\pi(t) + T^*_\alpha(t)\mu(t) = 0. \tag{5.10}$$

Set $\chi = (y, v, \pi)$, $\lambda = (q, \varrho, \mu)$ and

$$\mathcal{P} : Z^2 \to (Z^2)^*, \quad \mathcal{P} = \frac{1}{2}\begin{pmatrix} D^2\mathcal{L}_0 & 0 \\ 0 & I \end{pmatrix}, \tag{5.11}$$

where I is the identity in $L^2(0,1;\mathbb{R}^l)$. Using definitions (2.20), (3.1) and (5.11) we can write the stationarity conditions (5.8)–(5.10) in the form

$$\mathcal{P}\widetilde{\chi}_\delta + (b - e) + \mathcal{C}^*_\alpha\widetilde{\lambda}_\delta = 0. \tag{5.12}$$

LEMMA 5.1. *If* (A1)–(A3) *hold, then for any* $\delta \in U^\infty$ *there exists a unique solution* $\widetilde{\chi}_\delta = (\widetilde{y}_\delta, \widetilde{v}_\delta, \widetilde{\pi}_\delta) \in Z^\infty$ *of* $(\widetilde{\mathrm{LO}})_\delta$ *and a unique associated Lagrange multiplier* $\widetilde{\lambda}_\delta = (\widetilde{q}_\delta, \widetilde{\varrho}_\delta, \widetilde{\mu}_\delta) \in Y^\infty$. *The pair* $(\widetilde{\chi}_\delta, \widetilde{\lambda}_\delta)$ *is the only stationary point of* $(\widetilde{\mathrm{LO}})_\delta$.

Proof. Note that in view of (2.10), (3.17) and (5.11) the following coercivity condition holds:

$$(\chi, \mathcal{P}\chi) \ge \widetilde{\gamma}\|\chi\|^2_{Z^2}, \quad \text{where } \widetilde{\gamma} = \min\{\gamma, 1\}, \tag{5.13}$$

for all $\chi = (y, v, \pi)$ such that (y, v) satisfies (3.18).

Hence the existence of a unique solution follows in a standard way. If we use (A1), (A2) and (5.13), then regularity of the solution, as well as existence, uniqueness and regularity of Lagrange multipliers can be easily deduced from the results of Hager and Mitter (see [20, 19]), where a more general problem including pure state constraints is considered. So we confine ourselves to proving uniqueness of the stationary points. To this end, it is enough to show that each stationary point corresponds to the minimum. Using definitions (5.1) and (5.11) we obtain

$$\widetilde{\mathcal{I}}(\chi,\delta)-\widetilde{\mathcal{I}}(\widetilde{\chi}_\delta,\delta)=(\mathcal{P}\widetilde{\chi}_\delta+(b-e),\chi-\widetilde{\chi}_\delta)+\tfrac{1}{2}(\chi-\widetilde{\chi}_\delta,\mathcal{P}(\chi-\widetilde{\chi}_\delta)). \tag{5.14}$$

Note that, if $\chi=(y,v,\pi)$ is feasible for $(\widetilde{\mathrm{LO}})_\delta$, then

$$(\mathcal{C}_\alpha^*\widetilde{\lambda}_\delta,\chi-\widetilde{\chi}_\delta)=(\widetilde{\lambda}_\delta,\mathcal{C}_\alpha(\chi-\widetilde{\chi}_\delta))\le 0, \tag{5.15}$$

and moreover, the pair $(y-\widetilde{y}_\delta, v-\widetilde{v}_\delta)$ satisfies (3.18). Hence using (5.12)–(5.15), we obtain

$$\begin{aligned}\widetilde{\mathcal{I}}(\chi,\delta)-\widetilde{\mathcal{I}}(\widetilde{\chi}_\delta,\delta)&=(\mathcal{P}\widetilde{\chi}_\delta+(b-e),\chi-\widetilde{\chi}_\delta)+\tfrac{1}{2}(\chi-\chi_\delta,\mathcal{P}(\chi-\widetilde{\chi}_\delta))\\&\ge(\mathcal{P}\widetilde{\chi}_\delta+(b-e)+\mathcal{C}_\alpha^*\widetilde{\lambda}_\delta,\chi-\widetilde{\chi}_\delta)+\tfrac{1}{2}(\chi-\widetilde{\chi}_\delta,\mathcal{P}(\chi-\widetilde{\chi}_\delta))\\&\ge\tfrac{1}{2}\widetilde{\gamma}(\|\chi-\widetilde{\chi}_\delta\|_{Z^2})\end{aligned}$$

for any feasible χ, which shows that $\widetilde{\chi}_\delta$ is a unique minimizer of $(\widetilde{\mathrm{LO}})_\delta$ and completes the proof of the lemma. ∎

Lemma 5.2. *If* (A1)–(A3) *hold, then there exists a constant* $l>0$ *such that*

$$\|\widetilde{\chi}_{\delta_1}-\widetilde{\chi}_{\delta_2}\|_{Z^s},\ \|\widetilde{\lambda}_{\delta_1}-\widetilde{\lambda}_{\delta_2}\|_{Y^s}\le l\|\delta_1-\delta_2\|_{U^s}\quad \textit{for all } \delta_1,\delta_2\in U^\infty \textit{ and all } s\in[2,\infty]. \tag{5.16}$$

Proof. First we show that (5.16) holds for $s=2$. In view of Lemma 2.3, we can introduce the following new variable in $(\widetilde{\mathrm{LO}})_\delta$:

$$\phi=\chi+\mathcal{J}\mathcal{C}_\alpha^*(\mathcal{C}_\alpha\mathcal{J}\mathcal{C}_\alpha^*)^{-1}(a-d), \tag{5.17}$$

where $\mathcal{J}:(Z^2)^*\to Z^2$ is the canonical isomorphism, and in view of Lemma 2.3, $(\mathcal{C}_\alpha\mathcal{J}\mathcal{C}_\alpha^*)^{-1}:Y^2\to Y^2$ exists, as in (3.13). Note that, as functions of the new variables, the constraints (5.2)–(5.4) become *independent* of δ, and for any feasible ϕ conditions (3.18) are satisfied. In terms of ϕ the cost functional $\widetilde{\mathcal{I}}$ takes the form

$$\tfrac{1}{2}(\phi,\mathcal{P}\phi)+(L\delta,\phi)+M, \tag{5.18}$$

where M is a constant independent of ϕ and $L\delta\in(Z^2)^*$ is a linear bounded function of δ. Let ϕ_{δ_i} be the solution of $(\widetilde{\mathrm{LO}})_{\delta_i}$, $i=1,2$, for $\delta_i\in U^\infty$. In view of (5.18), a well known optimality condition for convex optimization problems yields

$$(\mathcal{P}\phi_{\delta_i}+L\delta_i,\phi-\phi_{\delta_i})\ge 0\quad\text{for all feasible }\phi.$$

Since the feasible set is independent of δ, we can substitute $\phi=\phi_2$ and $\phi=\phi_1$, for $i=1$ and $i=2$, respectively. So, we get

$$(\mathcal{P}\phi_{\delta_1}+L\delta_1,\phi_{\delta_2}-\phi_{\delta_1})\ge 0,\quad(\mathcal{P}\phi_{\delta_2}+L\delta_2,\phi_{\delta_1}-\phi_{\delta_2})\ge 0.$$

Adding these inequalities and using (5.13) we obtain

$$\begin{aligned}\widetilde{\gamma}\|\phi_{\delta_2}-\phi_{\delta_1}\|_{Z^2}^2&\le(\phi_{\delta_2}-\phi_{\delta_1},\mathcal{P}(\phi_{\delta_2}-\phi_{\delta_1}))\\&\le(L(\delta_1-\delta_2),\phi_{\delta_2}-\phi_{\delta_1})\le c\|\delta_1-\delta_2\|_{U^2}\|\phi_{\delta_2}-\phi_{\delta_1}\|_{Z^2},\end{aligned}$$

i.e.,

$$\|\phi_{\delta_2}-\phi_{\delta_1}\|_{Z^2}\le c\|\delta_1-\delta_2\|_{U^2}\quad\text{for all }\delta_1,\delta_2\in U^\infty.$$

This estimate together with (5.17) yields

$$\|\widetilde{\chi}_{\delta_2}-\widetilde{\chi}_{\delta_1}\|_{Z^2}\le c\|\delta_1-\delta_2\|_{U^2}\quad\text{for all }\delta_1,\delta_2\in U^\infty. \tag{5.19}$$

Finally, from (5.12) we get

$$\widetilde{\lambda}_{\delta_2} - \widetilde{\lambda}_{\delta_1} = -(\mathcal{C}_\alpha \mathcal{C}_\alpha^*)^{-1} \mathcal{C}_\alpha \left\{ \mathcal{P}(\widetilde{\chi}_{\delta_2} - \widetilde{\chi}_{\delta_1}) - (e_2 - e_1) \right\},$$

which, in view of (5.19), yields

$$\|\widetilde{\lambda}_{\delta_2} - \widetilde{\lambda}_{\delta_1}\|_{Z^2} \le c\|\delta_1 - \delta_2\|_{U^2} \quad \text{for all } \delta_1, \delta_2 \in U^\infty, \tag{5.20}$$

and completes the proof of (5.16) for $s = 2$.

Note that (5.19) and (5.20) mean, in particular, that

$$\|\widetilde{y}_{\delta_2} - \widetilde{y}_{\delta_1}\|_{1,2}, \|\widetilde{q}_{\delta_2} - \widetilde{q}_{\delta_1}\|_{1,2} \le c\|\delta_1 - \delta_2\|_{U^2} \quad \text{for all } \delta_1, \delta_2 \in U^\infty,$$

which implies

$$\|\widetilde{y}_{\delta_2} - \widetilde{y}_{\delta_1}\|_\infty, \|\widetilde{q}_{\delta_2} - \widetilde{q}_{\delta_1}\|_\infty \le c\|\delta_1 - \delta_2\|_{U^2} \quad \text{for all } \delta_1, \delta_2 \in U^\infty. \tag{5.21}$$

To show (5.16) for $s \in (2, \infty]$, notice that, pointwise for almost all $t \in [0,1]$, equations (5.9) and (5.10), together with (5.6) and (5.7), can be viewed as the optimality system for the following parametric quadratic program:

$(\mathrm{QP})_{\varpi(t)}$ Minimize

$$\frac{1}{2}[v^*, \pi^*] Q(t) \begin{bmatrix} v \\ \pi \end{bmatrix} + \langle Q_{21}(t) y(t) + B(t)^* q(t) + b^4(t) - e^4(t), v \rangle$$

subject to

$$\langle \varTheta^i(t), v \rangle + \langle T_\alpha^i(t), \pi \rangle + \langle \varUpsilon^i(t), y(t) \rangle + (a^3)^i(t) - (d^3)^i(t) \begin{cases} = 0 & \text{if } i \in J_\alpha(t), \\ \le 0 & \text{if } i \in I_\alpha(t) \setminus J_\alpha(t), \\ \text{free} & \text{if } i \notin I_\alpha(t), \end{cases}$$

where

$$Q(t) = \begin{bmatrix} Q_{22}(t) & 0 \\ 0 & I \end{bmatrix}.$$

Here $(v, \pi) \in \mathbb{R}^{m+l}$ is the argument and $\varpi(t) = (y(t), q(t), e^4(t), d^3(t)) \in \mathbb{R}^{2n+m+l}$ is treated as the parameter, while $(b^4(t), a^3(t)) \in \mathbb{R}^{m+l}$ is a fixed element. The vector $\mu(t) \in \mathbb{R}^l$, given in (5.7), is the associated Lagrange multiplier. Hence $(\widetilde{v}_\delta(t), \widetilde{\pi}_\delta(t))$ and $\widetilde{\mu}_\delta(t)$ can be treated, respectively, as the solution and Lagrange multiplier of $(\mathrm{QP})_{\varpi(t)}$ for $\varpi_\delta(t) = (\widetilde{y}_\delta(t), \widetilde{q}_\delta(t), e^4(t), d^3(t))$, where e^4 and d^3 are the appropriate components of δ. Note that, in view of (3.30), we have

$$[v^*, \pi^*] Q(t) \begin{bmatrix} v \\ \pi \end{bmatrix} \ge \min\{\gamma, 1\}(|v|^2 + |\pi|^2) \tag{5.22}$$

for all $v \in \{\mathbb{R}^m \mid \langle \varTheta^i(t), v \rangle = 0 \ \forall i \in J_\alpha(t)\}$ and all $\pi \in \mathbb{R}^l$. The Lipschitz stability property of parametric quadratic programs was analyzed in [19]. By linear independence condition (2.10) and by (5.22), the assumptions of Theorem 3.1 in [19] are satisfied, and by that theorem, there exists a constant $k > 0$, depending only on β and γ, such that

$$\begin{aligned} &|\widetilde{v}_{\delta_2}(t) - \widetilde{v}_{\delta_1}(t)|, |\widetilde{\mu}_{\delta_2}(t) - \widetilde{\mu}_{\delta_1}(t)| \\ &\quad \le k(|\widetilde{y}_{\delta_2}(t) - \widetilde{y}_{\delta_1}(t)| + |\widetilde{q}_{\delta_2}(t) - \widetilde{q}_{\delta_1}(t)| + |e_2^4(t) - e_1^4(t)| + |d_2^3(t) - d_1^3(t)|) \end{aligned} \tag{5.23}$$

for all $\delta_1, \delta_2 \in U^\infty$. Note that $\|\cdot\|_{U^2} \le \|\cdot\|_{U^s}$ for $s \in (2, \infty]$. Hence, from (5.21) and (5.23) we obtain

$$|\widetilde{v}_{\delta_2}(t) - \widetilde{v}_{\delta_1}(t)|, |\widetilde{\mu}_{\delta_2}(t) - \widetilde{\mu}_{\delta_1}(t)| \le k(\|\delta_2(t) - \delta_1(t)\|_{U^s} + |e_2^4(t) - e_1^4(t)| + |d_2^3(t) - d_1^3(t)|),$$

and simple calculations show

$$\|\widetilde{v}_{\delta_2} - \widetilde{v}_{\delta_1}\|_s, \|\widetilde{\mu}_{\delta_2} - \widetilde{\mu}_{\delta_1}\|_s \le k\|\delta_2(t) - \delta_1(t)\|_{U^s} \quad \text{for } s \in (2, \infty]. \tag{5.24}$$

Estimates (5.24), together with the state and adjoint equations (5.5) and (5.8), show that (5.16) holds. ■

The following lemma establishes the relation between the stationary points of $(\widetilde{\mathrm{LO}})_\delta$ and those of $(\mathrm{LO})_\delta$.

LEMMA 5.3. *If* (A1)–(A3) *hold, then there exist constants* $\varsigma_1, \varsigma_2 > 0$ *such that for each* $\delta \in \mathcal{B}^{U^\infty}_{\varsigma_1}(0)$ *we have*

$$(\widetilde{y}_\delta, \widetilde{v}_\delta, \widetilde{\pi}_\delta) = (y_\delta, v_\delta, 0), \quad (\widetilde{q}_\delta, \widetilde{\varrho}_\delta, \widetilde{\mu}_\delta) = (q_\delta, \varrho_\delta, \mu_\delta), \tag{5.25}$$

where $(y_\delta, v_\delta, q_\delta, \varrho_\delta, \mu_\delta)$ *is the unique stationary point in* $\mathcal{B}^{W^\infty}_{\varsigma_2}(\eta_0)$ *of* $(\mathrm{LO})_\delta$.

Proof. Denote by $(\overline{\mathrm{LO}})_\delta$ another modification of $(\mathrm{LO})_\delta$, where inequality constraints (4.22) are changed as follows:

$$\langle \Upsilon^i(t), y(t)\rangle + \langle \Theta^i(t), v(t)\rangle + (a^3)^i(t) - (d^3)^i(t) \begin{cases} = 0 & \text{if } i \in J_\alpha(t), \\ \le 0 & \text{if } i \notin J_\alpha(t). \end{cases} \tag{5.26}$$

Clearly, by (4.19), for $\delta = 0$ we have

$$(\widetilde{y}_0, \widetilde{v}_0, \widetilde{\pi}_0, \widetilde{q}_0, \widetilde{\varrho}_0, \widetilde{\mu}_0) = (x_0, u_0, 0, p_0, \varrho_0, \nu_0). \tag{5.27}$$

By Lemma 5.2 we can choose $\varsigma_1 > 0$ so small that

$$\langle \Upsilon^i(t), \widetilde{y}_\delta(t) - \widetilde{y}_0\rangle + \langle \Theta^i(t), \widetilde{v}_\delta(t) - \widetilde{v}_0\rangle - (d^3)^i(t) \le \alpha \tag{5.28}$$

for all $\delta \in \mathcal{B}^{U^\infty}_{\varsigma_1}(0)$, where $\alpha > 0$ is given in (A1). In view of (2.7), (4.18) and (5.27), inequality (5.28) implies that

$$\langle \Upsilon^i(t), \widetilde{y}_\delta(t)\rangle + \langle \Theta^i(t), \widetilde{v}_\delta(t)\rangle + (a^3)^i(t) - (d^3)^i(t) \le 0, \tag{5.29}$$
$$\text{for all } i \notin I_\alpha(t) \text{ and for a.a. } t \in [0,1].$$

On the other hand, by definition (2.9), $T_\alpha(t) = 0$ for $t \in I_\alpha(t)$. So (5.4) together with (5.29) shows that $(\widetilde{y}_\delta, \widetilde{v}_\delta)$ is feasible for $(\overline{\mathrm{LO}})_\delta$. We now prove that $(\widetilde{y}_\delta, \widetilde{v}_\delta)$ is the minimizer of $(\overline{\mathrm{LO}})_\delta$. Suppose the contrary, i.e., that there exists a pair $(\overline{y}_\delta, \overline{v}_\delta)$ feasible for $(\overline{\mathrm{LO}})_\delta$ such that $\mathcal{I}(\overline{y}_\delta, \overline{u}_\delta, \delta) < \mathcal{I}(\widetilde{y}_\delta, \widetilde{u}_\delta, \delta)$. Then we would have

$$\widetilde{\mathcal{I}}(\overline{y}_\delta, \overline{u}_\delta, 0, \delta) = \mathcal{I}(\overline{y}_\delta, \overline{u}_\delta, \delta) < \mathcal{I}(\widetilde{y}_\delta, \widetilde{u}_\delta, \delta) \le \widetilde{\mathcal{I}}(\widetilde{y}_\delta, \widetilde{u}_\delta, \widetilde{\pi}_\delta, \delta).$$

Since $(\overline{y}_\delta, \overline{u}_\delta, 0)$ is feasible for $(\widetilde{\mathrm{LO}})_\delta$, the above inequality contradicts optimality of $(\widetilde{y}_\delta, \widetilde{u}_\delta, \widetilde{\pi}_\delta)$ and shows that $(\widetilde{y}_\delta, \widetilde{u}_\delta)$ is the minimizer of $(\overline{\mathrm{LO}})_\delta$. Clearly, $(\widetilde{q}_\delta, \widetilde{\varrho}_\delta, \widetilde{\mu}_\delta)$ is the associated Lagrange multiplier. In view of (3.15) and (5.16), we can shrink $\varsigma > 0$ so that $\widetilde{\mu}^i_\delta > 0$ for $i \in J_\alpha(t)$, i.e., $(\widetilde{y}_\delta, \widetilde{v}_\delta, \widetilde{q}_\delta, \widetilde{\varrho}_\delta, \widetilde{\mu}_\delta)$ is a stationary point of $(\mathrm{LO})_\delta$.

To show local uniqueness of the stationary points of $(\mathrm{LO})_\delta$, choose $\varsigma_1, \varsigma_2 > 0$ so small that

$$\langle \Upsilon^i(t), y\rangle + \langle \Theta^i(t), v\rangle + (a^3)^i(t) - (d^3)^i < 0 \tag{5.30}$$
$$\text{for all } i \notin I_\alpha(t),\ |(d^3)^i| \le \varsigma_1,\ |y| \le \varsigma_2,\ |v| \le \varsigma_2 \text{ and for a.a. } t \in [0,1].$$

Let $\delta \in \mathcal{B}^{U^\infty}_{\varsigma_1}(0)$ and let $\eta_\delta := (y_\delta, v_\delta, q_\delta, \varrho_\delta, \mu_\delta) \in \mathcal{B}^{W^\infty}_{\varsigma_2}(\eta_0)$ be a stationary point of $(\mathrm{LO})_\delta$. In view of (5.30), the complementarity condition implies $\mu^i_\delta(t) = 0$ for $i \notin I_\alpha(t)$.

Hence it is easy to see that $(y_\delta, v_\delta, 0, q_\delta, \varrho_\delta, \mu_\delta)$ satisfies (5.5)–(5.10), i.e., it is a stationary point of $(\widetilde{\mathrm{LO}})_\delta$, which is unique. This implies uniqueness of η_δ and completes the proof of the lemma. ■

Lemmas 5.1–5.3 yield

Proposition 5.4. *If* (A1)–(A3) *hold, then there exist constants* $\varsigma_1, \varsigma_2, l > 0$ *such that for each* $\delta \in \mathcal{B}_{\varsigma_1}^{U^\infty}(0)$ *there is a unique stationary point* $(y_\delta, v_\delta, q_\delta, \varrho_\delta, \mu_\delta) \in \mathcal{B}_{\varsigma_2}^{W^\infty}(\eta_0)$ *of* $(\mathrm{LO})_\delta$ *and*

$$\begin{aligned}(5.31)\qquad &\|y_{\delta_2} - y_{\delta_1}\|_{1,s}, \|v_{\delta_2} - v_{\delta_1}\|_s, \|q_{\delta_2} - q_{\delta_1}\|_{1,s}, |\varrho_{\delta_2} - \varrho_{\delta_1}|, \|\mu_{\delta_2} - \mu_{\delta_1}\|_s \\ &\qquad \le l\|\delta_2 - \delta_1\|_{U^s} \quad \textit{for all } \delta_1, \delta_2 \in \mathcal{B}_{\varsigma_1}^{U^\infty}(0) \textit{ and all } s \in [2, \infty].\end{aligned}$$

Bibliographical note. A proof of the stability result based on the analysis of the modified problem $(\widetilde{\mathrm{LO}})_\delta$ was first given in [35]. Proposition 5.4, for $s = 2$, was proved in [34] and for $s = \infty$ in [13].

6. Differentiability of solutions to accessory problems

We are now going to investigate differentiability properties of the solutions to $(\mathrm{O})_h$. Since we will use Theorem 4.3, in this section we concentrate on differentiability of solutions of the linearized generalized equation (4.14)–(4.17), i.e., of the stationary points of the accessory problem $(\mathrm{LO})_\delta$. We show that the stationary points of $(\mathrm{LO})_\delta$ are Bouligand differentiable functions of δ. The proof is in two steps. First we show directional differentiability and characterize the directional differentials. Then we prove that the differential is uniform with respect to the direction, so it is the B-differential.

Lemma 6.1. *Let* (A1)–(A3) *hold and let* $\varsigma_1, \varsigma_2 > 0$ *be as in Proposition* 5.4. *Then the mapping*

$$\eta_\delta := (y_\delta, v_\delta, q_\delta, \varrho_\delta, \mu_\delta) : \mathcal{B}_{\varsigma_1}^{U^\infty}(0) \to X^2 \times Y^2$$

given by the stationary points in $\mathcal{B}_{\varsigma_2}^{W^\infty}(\eta_0)$ *of* $(\mathrm{LO})_\delta$ *is directionally differentiable. The directional differential at* $\delta = 0$ *in a direction* $\pi = (\pi^1, \pi^2, \pi^3, \pi^4, \pi^5, \pi^6, \pi^7) \in U^\infty$ *is given by the stationary point of the following linear-quadratic optimal control problem*:

$(\mathrm{LQ})_\pi$ *Find* $(z_\pi, w_\pi) \in X^\infty$ *such that*

$$\mathcal{J}(z_\pi, w_\pi, \pi) = \min\Big\{\mathcal{J}(z, w, \pi) = \tfrac{1}{2}((z, w), \mathcal{L}_0(z, w)) + \int_0^1 [\langle \pi^4(t), z(t)\rangle + \langle \pi^7(t), w(t)\rangle]\, dt + \langle \pi^5, y(0)\rangle + \langle \pi^6, y(1)\rangle\Big\},$$

subject to

$$\begin{aligned}&\dot z(t) - A(t)z(t)) - B(t)w(t) - \pi^1(t) = 0,\\ &\Xi_0 z(0) + \Xi_1 z(1) - \pi^2 = 0,\\ &\langle \Upsilon^i(t), z(t)\rangle + \langle \Theta^i(t), w(t)\rangle - \pi^3(t) \begin{cases} = 0 & \textit{if } i \in J_0(t),\\ \le 0 & \textit{if } i \in I_0(t) \setminus J_0(t),\\ \textit{free} & \textit{if } i \notin I_0(t).\end{cases}\end{aligned}$$

Proof. Choose $\pi \in U^\infty$ and let $\{\tau_k\} \downarrow 0$ be an arbitrary sequence of positive numbers tending to zero. Define $\delta_k = \tau_k \pi$ and let $\eta_k = (y_k, v_k, q_k, \varrho_k, \mu_k)$ be the stationary point of $(\mathrm{LO})_{\delta_k}$. By Proposition 5.4 we have

$$\|y_k - y_0\|_{1,\infty}, \|q_k - q_0\|_{1,\infty} \le l\|\tau_k\pi\|_{U^\infty}, \tag{6.1}$$

$$\|v_k - v_0\|_\infty, |\varrho_k - \varrho_0|, \|\mu_k - \mu_0\|_\infty \le l\|\tau_k\pi\|_{U^\infty}. \tag{6.2}$$

By (6.1),

$$\left\|\frac{y_k - y_0}{\tau_k}\right\|_{1,2}, \left\|\frac{q_k - q_0}{\tau_k}\right\|_{1,2} \le l\|\pi\|_{U^\infty},$$

whereas, by (6.2),

$$\left|\frac{\varrho_k - \varrho_0}{\tau_k}\right| \le l\|\pi\|_{U^\infty}.$$

Hence, there exists a subsequence, still denoted by $\{\tau_k\}$, and elements $z, r \in W^{1,2}(0,1;\mathbb{R}^n)$ as well as $\varpi \in \mathbb{R}^n$ such that

$$\left.\begin{array}{c} \dfrac{y_k - y_0}{\tau_k} \rightharpoonup z, \quad \dfrac{q_k - q_0}{\tau_k} \rightharpoonup r \text{ weakly in } W^{1,2}(0,1;\mathbb{R}^n) \\ \dfrac{\varrho_k - \varrho_0}{\tau_k} \to \varpi \end{array}\right\} \text{ as } \tau_k \to 0. \tag{6.3}$$

It is well known that the embedding $W^{1,2}(0,1;\mathbb{R}^n) \subset L^2(0,1;\mathbb{R}^n)$ is compact. So (6.3) implies

$$\frac{y_k - y_0}{\tau_k} \to z, \quad \frac{q_k - q_0}{\tau_k} \to r \quad \text{strongly in } L^2(0,1;\mathbb{R}^n).$$

Hence, in particular,

$$\frac{y_k(t) - y_0(t)}{\tau_k} \to z(t), \quad \frac{q_k(t) - q_0(t)}{\tau_k} \to r(t) \quad \text{for a.a. } t \in [0,1]. \tag{6.4}$$

For $\delta_k = \tau_k\pi$, equations (4.17) and (4.15) can be rewritten in the form

$$\begin{cases} Q_{22}(t)v_k(t) + Q_{21}(t)y_k(t) + B(t)^*(t)q_k(t) + \Theta(t)^*\mu_k(t) + b^4(t) - \tau_k\pi^7(t) = 0, \\ (\langle\Theta^i(t), v_k(t)\rangle + \langle\Upsilon^i(t), y_k(t)\rangle + (a^3)^i(t) - (\tau_k\pi^3)^i(t))\mu_k^i(t) = 0, \\ \langle\Theta^i(t), v_k(t)\rangle + \langle\Upsilon^i(t), y_k(t)\rangle + (a^3)^i(t) - (\tau_k\pi^3)^i(t) \le 0, \\ \mu_k^i(t) \ge 0, \end{cases} \tag{6.5}$$

for all $i \in I$ and a.a. $t \in [0,1]$, where the inclusion (4.15) is substituted by the equivalent pointwise (KKT)-conditions.

Note that by (6.2) we have

$$\left|\frac{v_k(t) - v_0(t)}{\tau_k}\right|, \left|\frac{\mu_k(t) - \mu_0(t)}{\tau_k}\right| \le l\|\pi\|_{U^\infty}.$$

So, for a subsequence

$$\frac{v_k(t) - v_0(t)}{\tau_k} \to w(t), \quad \frac{\mu_k(t) - \mu_0(t)}{\tau_k} \to \kappa(t) \quad \text{as } \tau_k \to 0. \tag{6.6}$$

Let us take the difference of (6.5) evaluated at $\tau_k\pi$ and at 0, and divide by τ_k. Passing to the limit as $\tau_k \to 0$ and using (6.4) and (6.6) as well as definitions (2.7) and (3.15), we get the following system:

$$(6.7)\qquad \begin{cases} Q_{22}(t)w(t) + Q_{21}(t)z(t) + B(t)^*(t)r(t) + \Theta(t)^*\kappa(t) - \pi^7(t) = 0, \\ (\langle \Theta^i(t), w(t)\rangle + \langle \Upsilon^i(t), z(t)\rangle - (\pi^3)^i(t))\kappa^i(t) = 0, \\ \langle \Theta^i(t), w(t)\rangle + \langle \Upsilon^i(t), z(t)\rangle - (\pi^3)^i(t) \begin{cases} = 0 & \text{for } i \in J_0(t), \\ \le 0 & \text{for } i \in I_0(t) \setminus J_0(t), \end{cases} \\ \kappa^i(t) \begin{cases} \ge 0 & \text{for } i \in I_0(t) \setminus J_0(t), \\ = 0 & \text{for } i \notin I_0(t). \end{cases} \end{cases}$$

An inspection of (6.7) shows that the pair $(w(t), \kappa(t)) \in \mathbb{R}^{n+l}$ can be treated as a stationary point of the following quadratic program $(\mathrm{M})_{\phi(t)}$ depending on the vector parameter $\phi(t) = (z(t), r(t), \pi^3(t), \pi^7(t)) \in \mathbb{R}^{2n+l+m}$:

$(\mathrm{M})_{\phi(t)}$ Minimize $\frac{1}{2}\langle w, Q_{22}(t)w\rangle + \langle Q_{21}(t)z(t) + B(t)^*r(t) - \pi^7(t), w\rangle$ subject to

$$\langle \Theta^i(t), w\rangle + \langle \Upsilon^i(t), z(t)\rangle - (\pi^3)^i(t) \begin{cases} = 0 & \text{for } i \in J_0(t), \\ \le 0 & \text{for } i \in I_0(t) \setminus J_0(t), \\ \text{free} & \text{for } i \notin I_0(t). \end{cases}$$

In view of (2.15) and (3.30), problem $(\mathrm{M})_{\phi(t)}$ has a unique stationary point $(w(t), \kappa(t))$, where $w(t)$ is the solution and $\kappa(t)$ the associated Lagrange multiplier. This shows that convergence in (6.6) takes place for the whole sequence.

By the Lebesgue dominated convergence theorem, the pointwise convergence (6.6), together with the estimate (6.2), implies

$$(6.8)\qquad \frac{v_k - v_0}{\tau_k} \to w, \quad \frac{\mu_k - \mu_0}{\tau_k} \to \kappa$$

strongly in $L^2(0,1;\mathbb{R}^m)$ and $L^2(0,1;\mathbb{R}^l)$, respectively.

Using (6.3) and (6.8) in the state and adjoint equations (4.14) and (4.16), we find that

$$(6.9)\qquad \begin{cases} \dot{z} - Az - Bw - \pi^1 = 0, \\ \Xi_0 z(0) + \Xi_1 z(1) - \pi^2 = 0, \end{cases}$$

$$(6.10)\qquad \begin{cases} \dot{r} + A^*r + Q_{11}z + Q_{12}w + \Upsilon^*\kappa - \pi^4 = 0, \\ r(0) + \mathcal{R}_{00}z(0) + \mathcal{R}_{01}z(1) + \Xi_0\varpi - \pi^5 = 0, \\ -r(1) + \mathcal{R}_{10}z(0) + \mathcal{R}_{11}z(1) + \Xi_1\varpi - \pi^6 = 0. \end{cases}$$

Equations (6.9) and (6.10) together with (6.7) constitute an optimality system for $(\mathrm{LQ})_\pi$. Note that, as in the case of $(\widetilde{\mathrm{LO}})_\delta$, conditions (A1)–(A3) ensure that, for any $\pi \in U^\infty$, $(\mathrm{LQ})_\pi$ has a unique stationary point, which corresponds to the solution and Lagrange multiplier. Hence the element $(z, w, r, \varpi, \kappa)$ is defined uniquely, i.e., the convergence in (6.3) and (6.6) holds for the whole sequence $\{\tau_k\}$. ■

Note that, using the same argument as in the proof of Lemma 5.2, we find that there exists a constant $l > 0$ such that

$$\|(z_{\pi_1} - z_{\pi_2}, w_{\pi_1} - w_{\pi_2})\|_{X^s} \le l\|\pi_1 - \pi_2\|_{U^s},$$
$$\|(r_{\pi_1} - r_{\pi_2}, \varpi_{\pi_1} - \varpi_{\pi_2}, \kappa_{\pi_1} - \kappa_{\pi_2})\|_{Y^s} \le l\|\pi_1 - \pi_2\|_{U^s}$$

for all $\pi_1, \pi_2 \in U^\infty$ and all $s \in [2, \infty]$. Since $(z_0, w_0, q_0, \varpi_0, \kappa_0) = (0, 0, 0, 0, 0)$, we have in particular

$$\|(z_\pi, w_\pi)\|_{X^s}, \|(r_\pi, \varpi_\pi, \kappa_\pi)\|_{X^s} \le l\|\pi\|_{U^s} \quad \text{for all } \pi \in U^\infty. \tag{6.11}$$

PROPOSITION 6.2. *Let* (A1)–(A3) *be satisfied and let* $\varsigma_1, \varsigma_2 > 0$ *be as in Proposition* 5.3. *Then the mapping*

$$\eta_\delta := (y_\delta, v_\delta, q_\delta, \varrho_\delta, \mu_\delta) : \mathcal{B}^{U^\infty}_{\varsigma_1} \to X^s \times Y^s$$

given by the stationary point in $\mathcal{B}^{W^\infty}_{\varsigma_2}(\eta_0)$ *of* $(\mathrm{LO})_\delta$ *is Bouligand differentiable for any* $s \in [2, \infty)$. *The B-differential at* $\delta = 0$ *in a direction* $\pi = (\pi^1, \pi^2, \pi^3, \pi^4, \pi^5, \pi^6, \pi^7) \in U^\infty$ *is given by the stationary point of the linear-quadratic optimal control problem* $(\mathrm{LQ})_\pi$.

Clearly, the stationary point $(z_\pi, w_\pi, r_\pi, \varpi_\pi, \kappa_\pi)$, characterized by (6.9), (6.10) and (6.7), is a positively homogeneous function of the perturbation π. Hence, in view of Definition 4.2, to prove the proposition, it is enough to show that, for any $s \in [2, \infty)$ and any $\varepsilon > 0$ there exists $\varsigma > 0$ such that

$$\begin{cases} \|(y_\pi, v_\pi) - (y_0, v_0) - (z_\pi, w_\pi)\|_{X^s} \le \varepsilon\|\pi\|_{U^\infty}, \\ \|(q_\pi, \varrho_\pi, \mu_\pi) - (q_0, \varrho_0, \mu_0) - (r_\pi, \varpi_\pi, \kappa_\pi)\|_{Y^s} \le \varepsilon\|\pi\|_{U^\infty}, \end{cases} \tag{6.12}$$

for any $\pi \in \mathcal{B}^{U^\infty}_\varsigma(0)$. Let us subtract (4.14), (4.16), (4.17) evaluated at $\delta = \pi$ and at $\delta = 0$. We obtain

$$\begin{cases} (\dot y_\pi - \dot y_0) - A(y_\pi - y_0) - B(v_\pi - v_0) - \pi^1 = 0, \\ \Xi_0(y_\pi(0) - y_0(0)) + \Xi_1(y_\pi(1) - y_0(1)) - \pi^2 = 0, \end{cases} \tag{6.13}$$

$$\begin{cases} (\dot q_\pi - \dot q_0) + A^*(q_\pi - q_0) + Q_{11}(y_\pi - y_0) + Q_{12}(v_\pi - v_0) \\ \qquad + \Upsilon^*(\mu_\pi - \mu_0) - \pi^4 = 0, \\ (q_\pi(0) - q_0(0)) + \mathcal{R}_{00}(y_\pi(0) - y_0(0)) + \mathcal{R}_{01}(y_\pi(0) - y_0(0)) \\ \qquad + \Xi_0(\varrho_\pi - \varrho_0) - \pi^5 = 0, \\ -(q_\pi(1) - q_0(1)) + \mathcal{R}_{10}(y_\pi(1) - y_0(0)) + \mathcal{R}_{11}(y_\pi(0) - y_0(0)) \\ \qquad + \Xi_1(\varrho_\pi - \varrho_0) - \pi^6 = 0, \end{cases} \tag{6.14}$$

$$Q_{21}(y_\pi - y_0) + Q_{22}(v_\pi - v_0) + B^*(q_\pi - q_0) + \Theta^*(\mu_\pi - \mu_0) - \pi^7 = 0. \tag{6.15}$$

To analyze (4.15), for a fixed $\beta > 0$ define the sets

$$\begin{cases} \mathcal{M}^i_\beta = \{t \in [0,1] \mid \theta^i(x_0(t), u_0(t), h_0(t)) \\ \qquad\qquad = \langle \Upsilon^i(t), y_0(t)\rangle + \langle \Theta^i(t), v_0(t)\rangle + (a^3)^i(t) \in (0, \beta)\}, \\ \mathcal{N}^i_\beta = \{t \in [0,1] \mid \nu^i_0(t) = \mu^i_0(t) \in (0, \beta)\}, \\ \mathcal{M}_\beta = \bigcup_{i\in I} \left(\mathcal{M}^i_\beta \cup \mathcal{N}^i_\beta\right). \end{cases} \tag{6.16}$$

It follows from (5.30) and (6.16) that, for any $\beta > 0$, there exists $\varsigma(\beta) > 0$ such that, for any $\pi \in \mathcal{B}^{U^\infty}_{\varsigma(\beta)}(0)$ and for almost all $t \in [0,1] \setminus \mathcal{M}_\beta$ we have

$$\mu^i_\pi(t) > 0 \;\Rightarrow\; \langle \Upsilon^i(t), y_\pi(t)\rangle + \langle \Theta^i(t), v_\pi(t)\rangle + (a^3)^i(t) - (\pi^3)^i(t) = 0 \quad \text{for } i \in J_0(t),$$

$$\left.\begin{aligned} &(\langle \Upsilon^i(t), y_\pi(t)\rangle + \langle \Theta^i(t), v_\pi(t)\rangle + (a^3)^i(t) - (\pi^3)^i(t))\mu^i_\pi(t) = 0 \text{ and} \\ &\langle \Upsilon^i(t), y_\pi(t)\rangle + \langle \Theta^i(t), v_\pi(t)\rangle + (a^3)^i(t) - (\pi^3)^i(t) \le 0,\; \mu^i_\pi(t) \ge 0 \end{aligned}\right\} \quad \text{for } i \in I_0(t) \setminus J_0(t),$$

$$\langle \Upsilon^i(t), y_\pi(t)\rangle + \langle \Theta^i(t), v_\pi(t)\rangle + (a^3)^i(t) - (\pi^3)^i(t) < 0 \;\Rightarrow\; \mu^i_\pi(t) = 0 \quad \text{for } i \notin I_0(t).$$

Hence, in view of (2.7) and (3.15), for $t \in [0,1] \setminus \mathcal{M}_\beta$, we obtain

$$
\text{(6.17)}\quad \begin{cases}
\langle \Upsilon^i(t), y_\pi(t) - y_0(t)\rangle + \langle \Theta^i(t), v_\pi(t) - v_0(t)\rangle \\
\qquad - (\pi^3)^i(t) = 0 & \text{for } i \in J_0(t), \\
\left.\begin{array}{l}
(\langle \Upsilon^i(t), y_\pi(t) - y_0(t)\rangle + \langle \Theta^i(t), v_\pi(t) - v_0(t)\rangle \\
\qquad - (\pi^3)^i(t))(\mu_\pi^i(t) - \mu_0^i(t)) = 0, \text{ and} \\
\langle \Upsilon^i(t), y_\pi(t) - y_0(t)\rangle + \langle \Theta^i(t), v_\pi(t) - v_0(t)\rangle \\
\qquad - (\pi^3)^i(t) \le 0 \\
\mu_\pi^i(t) - \mu_0(t) \ge 0,
\end{array}\right\} & \text{for } i \in I_0(t) \setminus J_0(t), \\
\mu_\pi^i(t) - \mu_0(t) = 0 & \text{for } i \notin I_0(t).
\end{cases}
$$

For $t \in \mathcal{M}_\beta$ we introduce the following new variables:

$$
\text{(6.18)}\quad \begin{array}{lll}
(\pi^4)'(t) = \pi^4(t) + (\Delta\pi^4)'(t), & \text{where} & (\Delta\pi^4)'(t) = -\Upsilon(t)^*(\mu_\pi(t) - \mu_0(t)), \\
(\pi^7)'(t) = \pi^7(t) + (\Delta\pi^7)'(t), & \text{where} & (\Delta\pi^7)'(t) = -\Theta(t)^*(\mu_\pi(t) - \mu_0(t)).
\end{array}
$$

Then the adjoint equation in (6.14) and equation (6.15) become

$$
\text{(6.19)}\quad \left.\begin{array}{l}
(\dot q_\pi - \dot q_0) + A^*(q_\pi - q_0) + Q_{11}(y_\pi - y_0) \\
\qquad + Q_{12}(v_\pi - v_0) - (\pi^4)' = 0, \\
Q_{21}(t)(y_\pi(t) - y_0(t)) + Q_{22}(t)(v_\pi(t) - v_0(t)) \\
\qquad + B(t)^*(q_\pi(t) - q_0(t)) - (\pi^7)'(t) = 0
\end{array}\right\} \quad \text{for } t \in \mathcal{M}_\beta.
$$

Let us introduce the following modification $(\overline{\mathrm{LQ}})_\pi^\beta$ of $(\mathrm{LQ})_\pi$, where the control constraints are void on the set $\mathcal{M}_\beta$:

$(\overline{\mathrm{LQ}})_\pi^\beta$ Minimize $\mathcal{J}(z, w, \pi)$ subject to

$$
\begin{array}{l}
\dot z(t) - A(t)z(t) - B(t)w(t) = \pi^1(t), \\
\Xi_0 z(0) + \Xi_1 z(1) = \pi^2, \\
\langle \Upsilon^i(t), z(t)\rangle + \langle \Theta^i(t), v(t)\rangle - (\pi^3)^i(t) \\
\qquad \begin{cases}
\left.\begin{array}{ll}
= 0 & \text{if } i \in J_0(t), \\
\le 0 & \text{if } i \in I_0(t) \setminus J_0(t), \\
\text{free} & \text{if } i \notin I_0(t),
\end{array}\right\} & \text{for } t \in [0,1] \setminus \mathcal{M}_\beta, \\
\text{free} & \text{for } t \in \mathcal{M}_\beta.
\end{cases}
\end{array}
$$

An inspection of (6.13)–(6.15), (6.17) and (6.19) shows that $(y_\pi - y_0, v_\pi - v_0, q_\pi - q_0, \varrho_\pi - \varrho_0, \mu_\pi - \mu_0)$ is a stationary point of $(\overline{\mathrm{LQ}})_{\pi'}^\beta$, where

$$
\pi' = \pi + \Delta\pi' \quad \text{with } \Delta\pi' = (0, 0, 0, (\Delta\pi^4)', 0, 0, (\Delta\pi^7)') \quad \text{and}
$$

$$
\text{(6.20)}\quad \begin{array}{l}
(\Delta\pi^4)'(t) = \begin{cases} 0 & \text{for } t \in [0,1] \setminus \mathcal{M}_\beta, \\ -\Upsilon(t)^*(\mu_\pi(t) - \mu_0(t)) & \text{for } t \in \mathcal{M}_\beta, \end{cases} \\
(\Delta\pi^7)'(t) = \begin{cases} 0 & \text{for } t \in [0,1] \setminus \mathcal{M}_\beta, \\ -\Theta(t)^*(\mu_\pi(t) - \mu_0(t)) & \text{for } t \in \mathcal{M}_\beta. \end{cases}
\end{array}
$$

Similarly, in view of (6.7), (6.9) and (6.10), $(z_\pi, w_\pi, r_\pi, \varpi_\pi, \kappa_\pi)$ can be interpreted as a stationary point of $(\overline{\mathrm{LQ}})_{\pi''}^\beta$, where

$$\pi'' = \pi + \Delta\pi'' \quad \text{with } \Delta\pi'' = (0,0,0,(\Delta\pi^4)'',0,0,(\Delta\pi^7)'') \quad \text{and}$$

$$(6.21) \qquad (\Delta\pi^4)''(t) = \begin{cases} 0 & \text{for } t \in [0,1] \setminus \mathcal{M}_\beta, \\ -\Upsilon(t)^*\kappa_\pi(t) & \text{for } t \in \mathcal{M}_\beta. \end{cases}$$
$$(\Delta\pi^7)''(t) = \begin{cases} 0 & \text{for } t \in [0,1] \setminus \mathcal{M}_\beta, \\ -\Theta(t)^*\kappa_\pi(t) & \text{for } t \in \mathcal{M}_\beta. \end{cases}$$

Note that, if $\beta \le \alpha$, where α is given in (A3), then (A3) implies that the quadratic term of the cost functional in $(\overline{\mathrm{LQ}})^\beta_\pi$ is coercive on the feasible set. Hence, as in $(\widetilde{\mathrm{LQ}})_\delta$, there exists a unique stationary point $(\overline{z}_\pi, \overline{w}_\pi, \overline{r}_\pi, \overline{\varpi}_\pi, \overline{\kappa}_\pi)$ of $(\overline{\mathrm{LQ}})^\beta_\pi$ and a constant $\bar{l} > 0$, independent of π and β, such that

$$(6.22) \qquad \begin{aligned} &\|(\overline{z}_{\pi_1}, \overline{w}_{\pi_1}) - (\overline{z}_{\pi_2}, \overline{w}_{\pi_2})\|_{X^s} \le \bar{l}\|\pi_1 - \pi_2\|_{U^s}, \\ &\|(\overline{r}_{\pi_1}, \overline{\varpi}_{\pi_1}, \overline{\kappa}_{\pi_1}) - (\overline{r}_{\pi_2}, \overline{\varpi}_{\pi_2}, \overline{\kappa}_{\pi_2})\|_{X^s} \le \bar{l}\|\pi_1 - \pi_2\|_{U^s}, \end{aligned}$$

for all $\pi_1, \pi_2 \in U^\infty$ and for all $s \in [2,\infty]$. Putting $\pi_1 = \pi'$, $\pi_2 = \pi''$ and using (6.22) together with (6.20) and (6.21) as well as (5.30) and (6.11), we obtain

$$(6.23) \qquad \begin{aligned} &\|(y_\pi - y_0 - z_\pi, v_\pi - v_0 - w_\pi)\|_{X^s} = \|(\overline{y}_{\pi'} - \overline{y}_{\pi''}, \overline{w}_{\pi'} - \overline{w}_{\pi''})\|_{X^s} \le \bar{l}\|\pi' - \pi''\|_{U^s} \\ &\le \bar{l}\Big[\int_{\mathcal{M}_\beta} (|\Theta(t)^*(\mu_\pi(t) - \mu_0(t) - \kappa_\pi(t))|^s + |\Upsilon(t)^*(\mu_\pi(t) - \mu_0(t) - \kappa_\pi(t))|^s)\,dt\Big]^{1/s} \\ &\le \bar{l}(\|\mu_\pi - \mu_0\|_\infty + \|\kappa_\pi\|_\infty)\Big[\int_{\mathcal{M}_\beta} (|\Theta(t)|^s + |\Upsilon(t)|^s)\ dt\Big]^{1/s} \\ &\le 2\bar{l}\|\pi\|_{U^\infty}\Big[\int_{\mathcal{M}_\beta} (|\Theta(t)|^s + |\Upsilon(t)|^s)\ dt\Big]^{1/s}. \end{aligned}$$

Similarly

$$(6.24) \qquad \begin{aligned} &\|(q_\pi - q_0 - r_\pi, \varrho_\pi - \varrho_0 - \varpi_\pi, \mu_\pi - \mu_0 - \kappa_\pi)\|_{Y^s} \\ &\qquad \le c\|\pi\|_{U^\infty}\Big[\int_{\mathcal{M}_\beta} (|\Theta(t)|^s + |\Upsilon(t)|^s)\,dt\Big]^{1/s}. \end{aligned}$$

Choose any $\varepsilon > 0$; since meas $\mathcal{M}_\beta \to 0$ as $\beta \to 0$, for any $s \in [2,\infty)$, we can find $\beta(\varepsilon, s) > 0$ and the corresponding $\overline{\varsigma} := \varsigma(\beta(\varepsilon, s))$ such that (6.12) holds for any $\pi \in \mathcal{B}^{U^\infty}_{\overline{\varsigma}}(0)$. ■

Bibliographical note. The proofs of Lemma 6.1 and of Proposition 6.2 can be found in [32] and in [37], respectively.

7. Lipschitz stability of solutions to nonlinear problems

Now we return to our nonlinear optimal control problem $(\mathrm{O})_h$. In this section the main stability result for this problem will be derived. In view of (5.30), for $s = \infty$, the abstract Theorem 4.1 implies:

PROPOSITION 7.1. *If* (A1)–(A3) *hold, then there exist constants* $\sigma_1, \sigma_2 > 0$ *and* $l > 0$ *such that for each* $h \in \mathcal{B}^H_{\sigma_1}(h_0)$ *there is a unique stationary point* $\zeta_h = (x_h, u_h, p_h, \varrho_h, \nu_h)$ *in* $\mathcal{B}^{Z^\infty}_{\sigma_2}(\zeta_0)$ *of* $(\mathrm{O})_h$ *and*

$$\begin{aligned}(7.1)\qquad &\|x_{h_2}-x_{h_1}\|_{1,\infty}, \|u_{h_2}-u_{h_1}\|_\infty, \|p_{h_2}-p_{h_1}\|_{1,\infty}, |\varrho_{h_2}-\varrho_{h_1}|, \|\nu_{h_2}-\nu_{h_1}\|_\infty \\ &\qquad \le l\|h_2-h_1\|_H \quad \textit{for all } h_1,h_2\in\mathcal{B}^H_{\sigma_1}(h_0).\end{aligned}$$

We will show that, for $\sigma_1>0$ sufficiently small, (x_h,u_h) is a solution of $(\mathrm{O})_h$ and (p_h,ϱ_h,ν_h) the associated Lagrange multiplier.

Let $(x_h,u_h,p_h,\varrho_h,\nu_h)$ be the stationary point of $(\mathrm{O})_h$, which exists for $h\in\mathcal{B}^H_{\sigma_1}(h_0)$ by Proposition 7.1. In the same way as in (2.20), introduce the mapping $\mathcal{C}^h_\alpha: Z^s\to Y^s$ defined by

$$\mathcal{C}^h_\alpha\begin{pmatrix} y\\ v\\ \vartheta\end{pmatrix} = \begin{pmatrix}\dot y - A^h y - B^h v\\ \Xi^h_0 y(0)+\Xi^h_1 y(1)\\ \Upsilon^h y+\Theta^h v + T^h_\alpha\vartheta\end{pmatrix}, \tag{7.2}$$

where the superscript "h" denotes that the respective elements are evaluated at (x_h,u_h,h), rather than at (x_0,u_0,h_0). Moreover, set

$$\mathcal{L}_h := \mathcal{L}(x_h,u_h,p_h,\varrho_h,\nu_h,h). \tag{7.3}$$

We will need the following auxiliary result.

Lemma 7.2. *If the assumptions of Proposition* 7.1 *hold*, *then for* $\sigma_1>0$ *sufficiently small*, *the mapping*

$$\mathcal{C}^h_\alpha: Z^s\to Y^s,\quad s\in[1,\infty],\quad \textit{is surjective for all } h\in\mathcal{B}^H_{\sigma_1}(h_0). \tag{7.4}$$

Proof. In view of (7.1), it follows from (2.20) and (7.2) that

$$\|\mathcal{C}^h_\alpha-\mathcal{C}_\alpha\|_{Z^s\to Y^s}\to 0\quad \text{as } \sigma_1\to 0. \tag{7.5}$$

Since $\mathcal{C}_\alpha: Z^s\to Y^s$ is surjective by Lemma 2.3, surjectivity of $\mathcal{C}^h_\alpha$ follows from (7.5) and from the Banach open mapping theorem, as in the proof of Lemma 9.3 below. ■

Lemma 7.3. *If the assumptions of Proposition* 7.1 *hold*, *then for* $\sigma_1>0$ *sufficiently small we have*

$$((y,v),D^2\mathcal{L}_h(y,v))\ge \frac{\gamma'}{2}(\|y\|^2_{1,2}+\|v\|^2_2) \tag{7.6}$$

for all $h\in\mathcal{B}^H_{\sigma_1}(h_0)$ *and all* $(y,v)\in X^2$ *such that*

$$(y,v)\in\ker\overline{\mathcal{C}}^h_\alpha, \tag{7.7}$$

where

$$\overline{\mathcal{C}}^h_\alpha\begin{pmatrix} y\\ v\end{pmatrix} = \begin{pmatrix}\dot y - A^h y - B^h v\\ \Xi^h_0 y(0)+\Xi^h_1 y(1)\\ \overline{\Upsilon}^h_\alpha y+\overline{\Theta}^h_\alpha v\end{pmatrix}. \tag{7.8}$$

Proof. In the same way as in (3.29) define

$$\mathcal{Q}^h_\alpha = \Gamma^h_\alpha(\mathcal{M}^h+Q^h_{22})\Gamma^h_\alpha, \tag{7.9}$$

where the mappings $\mathcal{M}^h$ and Γ^h_α as well as the subspace $(\mathcal{U}^h_\alpha)^s$ are defined as in (3.25)–(3.27), but evaluated at (ζ_h,h), rather than at the reference point (ζ_0,h_0). By the same construction as in (3.22) and (3.28) we find that (7.6) is equivalent to

$$(v,\mathcal{Q}^h_\alpha v)\ge\frac{\gamma}{2}\|\Gamma^h_\alpha v\|^2_2\quad\text{for all } v\in L^2(0,1;\mathbb{R}^m). \tag{7.10}$$

By (7.1) we have

$$\|\mathcal{Q}_\alpha^h - \mathcal{Q}_\alpha\|_{L^2\to L^2} \to 0, \quad \|\Gamma_\alpha^h - \Gamma_\alpha\|_{L^2\to L^2} \to 0 \quad \text{as } \|h-h_0\|_H \to 0.$$

Hence, in view of (3.28), we can find $\sigma_1 > 0$ such that (7.10) holds for all $h \in \mathcal{B}_{\sigma_1}^H(h_0)$. ∎

The following lemma shows that conditions (7.4) and (7.6) imply that (x_h, u_h) is an isolated local minimizer of order two for $(\mathrm{O})_h$.

LEMMA 7.4. *Let* (A1)–(A3) *hold. Then there exist* $\sigma_1, \sigma_2 > 0$ *such that for each* $h \in \mathcal{B}_{\sigma_1}^H(h_0)$ *and for each* $(x,u) \in \mathcal{B}_{\sigma_2}^{X^\infty}(x_h, u_h)$,

$$F(x,u,h) - F(x_h,u_h,h) \ge c(\|x-x_h\|_{1,2}^2 + \|u-u_h\|_2^2), \tag{7.11}$$

where $c > 0$ *is independent of* h.

Note that (7.11) holds for all feasible (x,u) in a X^∞-neighborhood of (x_h,u_h), whereas the quadratic term on the right-hand side of (7.11) is the square of the X^2-norm. These different norms reflect the phenomenon of *two-norm discrepancy.*

Proof of Lemma 7.4. Choose $\sigma_1 \le \alpha/(2l)$ such that the conclusions of Lemmas 7.2 and 7.3 hold. In view of (3.15) and (7.1), we have

$$\nu_h^i > \alpha/2 \quad \text{for } i \in J_\alpha(t), \tag{7.12}$$

Let $(x,u) \in X^\infty$ be feasible for $(\mathrm{O})_h$. By equality constraints and complementary slackness we have

$$\begin{cases} \mathcal{L}(x_h,u_h,p_h,\varrho_h,\nu_h,h) = F(x_h,u_h,h), \\ \mathcal{L}(x,u,p_h,\varrho_h,\nu_h,h) = F(x,u,h) + (\nu_h, \theta(x,u,h)). \end{cases} \tag{7.13}$$

Expanding $\mathcal{L}(\cdot,\cdot,p_h,\varrho_h,\nu_h,h)$ into Taylor series at (x_h,u_h) and using (7.13), as well as the stationarity condition (3.3) we obtain

$$\begin{aligned} F(x,u,h) - F(x_h,u_h,h) = &- (\nu_h, \theta(x,u,h)) \\ &+ ((x,u)-(x_h,u_h), D^2\mathcal{L}_h((x,u)-(x_h,u_h))) \\ &+ r_1((x,u),h) \end{aligned} \tag{7.14}$$

where

$$r_1((x,u),h) = \Big((x,u)-(x_h,u_h), \int_0^1 [D^2\mathcal{L}_s - D^2\mathcal{L}_h]\,ds\,((x,u)-(x_h,u_h))\Big)$$

and

$$D^2\mathcal{L}_s = D^2\mathcal{L}(x_h + s(x-x_h), u_h + s(u-u_h), p_h, \varrho_h, \nu_h, h), \quad s \in [0,1].$$

In view of (I) we have

$$\frac{|r_1((x,u),h)|}{\|(x,u)-(x_h,u_h)\|_{X^2}} \to 0 \quad \text{as } \|(x,u)-(x_h,u_h)\|_{X^\infty} \to 0, \tag{7.15}$$

uniformly with respect to $h \in \mathcal{B}_{\sigma_1}^H(h_0)$.

The first order expansion of the constraint functions (2.2)–(2.4) at (x_h,u_h) yields

$$\begin{cases} (\dot{x}-\dot{x}_h) - A^h(x-x_h) - B^h(u-u_h) = r_2^1((x,u),h), \\ \Xi_0^h(x(0)-x_h(0)) + \Xi_1^h(x(1)-x_h(1)) = r_2^2((x,u),h), \\ \Upsilon^h(x-x_h) + \Theta^h(u-u_h) = \theta(x,u,h) - \theta(x_h,u_h,h) + r_2^3((x,u),h), \end{cases} \tag{7.16}$$

where $r_2((x,u),h)=(r_2^1((x,u),h),r_2^2((x,u),h),r_2^3((x,u),h))\in Y^\infty$ satisfies

$$\frac{\|r_2((x,u),h)\|_{Y^2}}{\|(x,u)-(x_h,u_h)\|_{X^2}}\to 0 \quad \text{as } \|(x,u)-(x_h,u_h)\|_{X^\infty}\to 0. \tag{7.17}$$

Define the function $\overline{\theta}(x(t),u(t),h)$ as follows:

$$\overline{\theta}^i(x(t),u(t),h(t))=\begin{cases}\theta^i(x(t),u(t),h(t)) & \text{if } i\in J_\alpha(t),\\ 0 & \text{if } i\notin J_\alpha(t).\end{cases} \tag{7.18}$$

Since $\theta^i(x(t),u(t),h)\le 0$ and $\nu_h^i(t)\ge 0$, (7.12) implies

$$-(\nu_h,\theta(x,u,h))\ge -(\nu_h,\overline{\theta}(x,u,h))\ge \frac{\alpha}{2}\|\overline{\theta}(x,u,h)\|_1. \tag{7.19}$$

Consider the following equation:

$$\mathcal{C}_\alpha^h\begin{bmatrix} y\\ v\\ \vartheta\end{bmatrix}=\begin{bmatrix} r_2^1((x,u),h)\\ r_2^2((x,u),h)\\ r_2^3((x,u),h)\end{bmatrix}+\begin{bmatrix}0\\ 0\\ \overline{\theta}(x,u,h)\end{bmatrix}. \tag{7.20}$$

By Lemma 7.2 there exists a solution $(y,v,\vartheta)\in Z^2$ of (7.20) such that

$$\|(y,v,\vartheta)\|_{Z^2}\le k(\|r_2(x,u,h)\|_{Y^2}+\|\overline{\theta}(x,u,h)\|_2), \tag{7.21}$$

where $k>0$ is independent of $h\in\mathcal{B}_\sigma^H(h_0)$. Note that, since in view of (7.12),

$$\theta^i(x_h(t),u_h(t),h(t))=0 \quad \text{for } i\in J_\alpha(t), \tag{7.22}$$

equations (7.16) and (7.20) imply that the pair

$$(z,w):=(x-x_h,u-u_h)-(y,v) \tag{7.23}$$

satisfies condition (7.7). By (7.19) and (7.23), we obtain from (7.14)

$$\begin{aligned}F(x,u,h)-F(x_h,u_h,h)\ge \frac{\alpha}{2}\|\overline{\theta}(x,u,h)\|_1&+((z,w),D^2\mathcal{L}_h(z,w))\\ &+2((z,w),D^2\mathcal{L}_h(y,v))+((y,v),D^2\mathcal{L}_h(y,v))+r_1((x,u),h).\end{aligned}$$

Using Young's inequality as well as (7.6) and (7.21) we get

$$\begin{aligned}F(x,u,h)-F(x_h,u_h,h)&\ge \frac{\alpha}{2}\|\overline{\theta}(x,u,h)\|_1+\frac{\gamma'}{2}\|(z,w)\|_{X^2}^2-\frac{\gamma'}{4}\|(z,w)\|_{X^2}^2\\ &\quad-\frac{4}{\gamma'}\|D^2\mathcal{L}_h\|_{X^2\to X^2}^2\|(y,v)\|_{X^2}^2\\ &\quad-\|D^2\mathcal{L}_h\|_{X^2\to X^2}\|(y,v)\|_{X^2}^2+r_1((x,u),h)\\ &\ge \frac{\alpha}{2}\|\overline{\theta}(x,u,h)\|_1+\frac{\gamma'}{4}\|(z,w)\|_{X^2}^2-c\|\overline{\theta}(x,u,h)\|_2^2\\ &\quad-c\|r_2((x,u),h)\|_{Y^2}^2+r_1((x,u),h).\end{aligned}$$

We have

$$\|\overline{\theta}(x,u,h)\|_2^2\le\|\overline{\theta}(x,u,h)\|_1\|\overline{\theta}(x,u,h)\|_\infty,$$

while, in view of (7.18) and (7.22), $\|\overline{\theta}(x,u,h)\|_\infty\to 0$ as $\|(x,u)-(x_h,u_h)\|_\infty\to 0$. Hence we can choose $\sigma_2>0$ so small that

$$F(x,u,h)-F(x_h,u_h,h)\ge\frac{\alpha}{4}\|\overline{\theta}(x,u,h)\|_1+\frac{\gamma'}{4}\|(z,w)\|_{X^2}^2-c\|r_2((x,u),h)\|_{Y^2}^2+r_1((x,u),h).$$

Using again (7.21) as well as (7.15) and (7.17), we find that, for $\sigma_2 > 0$ sufficiently small,

$$F(x,u,h) - F(x_h,u_h,h) \geq c(\|(y,v)\|^2_{X^2} + \|(z,w)\|^2_{X^2}),$$

where $c > 0$ is independent of h, i.e., in view of (7.23), (7.11) holds. ∎

Proposition 7.1 together with Lemma 7.4 yields the following principal stability result for solutions to $(\mathrm{O})_h$.

THEOREM 7.5. *If conditions* (I)–(III) *and* (A1)–(A3) *hold, then there exist constants* $\sigma_1, \sigma_2, l > 0$ *such that for each* $h \in \mathcal{B}^H_{\sigma_1}(h_0)$ *there is a unique solution* (x_h, u_h) *in* $\mathcal{B}^{X^\infty}_{\sigma_2}(x_0,u_0)$ *of* $(\mathrm{O})_h$ *and a unique associated Lagrange multiplier* (p_h, ϱ_h, ν_h). *Moreover*,

$$\|x_{h_2} - x_{h_1}\|_{1,\infty}, \|u_{h_2} - u_{h_1}\|_\infty, \|p_{h_2} - p_{h_1}\|_{1,\infty}, |\varrho_{h_2} - \varrho_{h_1}|, \|\nu_{h_2} - \nu_{h_1}\|_\infty \\ \leq l\|h_2 - h_1\|_H \quad \textit{for all } h_1, h_2 \in \mathcal{B}^H_{\sigma_1}(h_0).$$

Bibliographical note. The proof of Lemma 7.4 is based on that in [13].

8. Differentiability of solutions to nonlinear problems

By Theorem 4.4, Proposition 6.2 and Theorem 7.5 we obtain the following basic sensitivity result for $(\mathrm{O})_h$.

THEOREM 8.1. *If* (I)–(III) *and* (A1)–(A3) *hold, then there exist* $\sigma_1, \sigma_2 > 0$ *such that the mappings*

$$(8.1) \qquad (x_h,u_h) : \mathcal{B}^H_{\sigma_1}(h_0) \to X^s, \quad (p_h,\varrho_h,\nu_h) : \mathcal{B}^H_{\sigma_1}(h_0) \to Y^s, \quad s \in [2,\infty),$$

given by the solutions in $\mathcal{B}^{X^\infty}_{\sigma_2}((x_0,u_0))$ *and Lagrange multipliers of* $(\mathrm{O})_h$, *are Bouligand differentiable functions of* h *and the B-differentials*, *in a given direction* $g \in H$, *evaluated at* h_0, *are given by the solution and Lagrange multiplier of the following linear-quadratic optimal control problem*:

$(\mathrm{L})_{h_0,g}$ *Find* $(y_{h_0,g}, v_{h_0,g})$ *such that*

$$\mathcal{K}(y_{h_0,g}, v_{h_0,g}, g) = \min\left\{\mathcal{K}(y,v,g) := \tfrac{1}{2}((y,v), D^2\mathcal{L}_0(y,v)) + (y, D^2_{xh}\mathcal{L}_0 g) + (v, D^2_{uh}\mathcal{L}_0 g)\right\}$$

subject to

$$\begin{aligned}
&\dot{y}(t) - D_x f_0(t)y(t) - D_u f_0(t)v(t) - D_h f_0(t)g(t) = 0,\\
&D_{x(0)}\xi_0 y(0) + D_{x(1)}\xi_0 y(1) + D_h\xi_0 g(0) = 0,\\
&\langle D_x\theta^i_0(t), y(t)\rangle + \langle D_u\theta^i_0(t), v(t)\rangle + \langle D_h\theta^i_0(t), g(t)\rangle \begin{cases} = 0 & \textit{for } i \in J_0(t),\\ \leq 0 & \textit{for } i \in I_0(t)\setminus J_0(t).\end{cases}
\end{aligned}$$

As noticed in Section 4, a Bouligand differential becomes Fréchet if it is linear and continuous. Hence, from the form of $(\mathrm{L})_{h_0,g}$, we obtain immediately

COROLLARY 8.2. *If*, *in addition to the assumptions of Theorem* 8.1, $I_0(t) = J_0(t)$ *for almost all* $t \in [0,1]$, *i.e.*, *if the pointwise strict complementarity holds*, *then the mappings* (8.1) *are Fréchet differentiable.*

The following simple example shows that, in general, the solutions to parametric optimal control problems are *not* differentiable in X^∞, so the result of Theorem 8.1 cannot be strengthened in this direction.

Let $H = \mathbb{R}$. In a neighborhood of the reference value $h_0 = 0$, consider the following parametric problem:

$(\mathrm{E})_h$ Find $(x_h, u_h) \in X^\infty := W^{1,\infty}(0,1;\mathbb{R}) \times L^\infty(0,1;\mathbb{R})$ such that

$$\frac{1}{2}\int_0^1 (x_h(t) - u_h(t))^2\, dt = \min \frac{1}{2}\int_0^1 (x(t) - u(t))^2\, dt$$

subject to

$$\dot{x}(t) - (2-h) = 0, \quad x(0) = 0, \quad u(t) - 1 \le 0.$$

It is easy to see that the solution to $(\mathrm{E})_h$ is given by

$$x_h(t) = (2-h)t, \quad u_h(t) = \begin{cases} (2-h)t & \text{for } t \in (0, 1/(2-h)), \\ 1 & \text{for } t \in (1/(2-h), 1). \end{cases}$$

So we have

$$\frac{du_0(t)}{dh} = \begin{cases} -t & \text{for } t \in (0,1/2), \\ 0 & \text{for } t \in (1/2, 1), \end{cases}$$

and a simple calculation yields

$$\operatorname*{ess\,sup}_{t\in[0,1]} \left| u_h(t) - u_0(t) - h\frac{du_0(t)}{dh} \right| = \frac{1}{2}|h|,$$

$$\left(\int_0^1 \left| u_h(t) - u_0(t) - h\frac{du_0(t)}{dh} \right|^s dt \right)^{1/s} = \frac{1}{2}(2-h)^{-1/s}(s+1)^{-1/s}|h|^{1+1/s}.$$

Thus, u_h is Fréchet differentiable at h_0 in X^s for any $s \in [2, \infty)$, but not in X^∞.

In sensitivity analysis of optimization problems an important role is played by the so-called optimal value function, which on $\mathcal{B}^H_{\sigma_1}(h_0)$ is defined as follows:

$$F^0(h) := F(x_h, u_h, h), \tag{8.2}$$

i.e., to each $h \in \mathcal{B}^H_{\sigma_1}(h_0)$, F^0 assigns the (local) optimal value of the cost functional. It has been known (see, e.g., [33]) that directional differentiability of the solutions implies the second order *directional* expansion of the value function. The following corollary of Theorem 8.1 shows that Bouligand differentiability of the solutions implies the second order expansion of F^0, *uniform* on a neighborhood of h_0.

Corollary 8.3. *If the assumptions of Theorem* 8.1 *hold*, *then for each* $h = h_0 + g \in \mathcal{B}^H_{\sigma_1}(h_0)$,

$$\begin{aligned} F^0(h) &= F^0(h_0) + (D_h\mathcal{L}_0, g) \\ &\quad + \frac{1}{2}\left((y_{h_0,g}, v_{h_0,g}, g), \begin{pmatrix} D^2_{xx}\mathcal{L}_0 & D^2_{xu}\mathcal{L}_0 & D^2_{xh}\mathcal{L}_0 \\ D^2_{ux}\mathcal{L}_0 & D^2_{uu}\mathcal{L}_0 & D^2_{uh}\mathcal{L}_0 \\ D^2_{hx}\mathcal{L}_0 & D^2_{hu}\mathcal{L}_0 & D^2_{hh}\mathcal{L}_0 \end{pmatrix} (y_{h_0,g}, v_{h_0,g}, g) \right) \\ &\quad + o(\|g\|^2_H), \end{aligned} \tag{8.3}$$

where $(y_{h_0,g}, v_{h_0,g})$ *is the B-differential of* (x_h, u_h) *at* h_0 *in direction* g, *i.e., it is given by the solution to* $(\mathrm{L})_{h_0,g}$.

Proof. In view of (7.13) and (8.2), we have

$$F^0(h) = \mathcal{L}(x_h, u_h, p_h, \varrho_h, \nu_h, h).$$

Using this equality and Theorem 8.1, for $h \in \mathcal{B}^H_{\sigma_1}(h_0)$, we get the following form of the B-differential of F^0:

$$D_h F^0(h)g = D_x\mathcal{L}_h y_{h,g} + D_u\mathcal{L}_h v_{h,g} + D_p\mathcal{L}_h q_{h,g} + D_\varrho\mathcal{L}_h \varrho_{h,g} + D_\nu\mathcal{L}_h \mu_{h,g} + D_h\mathcal{L}_h g,$$

where $(y_{h,g}, v_{h,g}, q_{h,g}, \varrho_{h,g}, \mu_{h,g})$ are given by the solution and Lagrange multipliers of $(\mathrm{L})_{h,g}$. By optimality conditions we have $D_x\mathcal{L}_h = 0$, $D_u\mathcal{L}_h = 0$. The partial derivatives of $\mathcal{L}_h$ with respect to the Lagrange multipliers give the corresponding constraints. So, for equality constraints we have $D_p\mathcal{L}_h = 0$, $D_\varrho\mathcal{L}_h = 0$. For inequality constraints, $\mu_{h,g}$ is the corresponding multiplier in $(\mathrm{L})_{h,g}$, so $\mu_{h,g}(t) = 0$ if $\theta_h(t) < 0$, i.e., $D_\nu\mathcal{L}_h\mu_{h,g} = (\theta_h, \mu_{h,g}) = 0$. Thus finally, we obtain

$$D_h F^0(h) = D_h\mathcal{L}_h. \tag{8.4}$$

Certainly, we have

$$F^0(h) = F^0(h_0) + \int_0^1 D_h F^0(h_\alpha) g\, d\alpha, \tag{8.5}$$

where $h_\alpha = h_0 + \alpha\Delta h$. Using (8.4) and Theorem 8.1 we obtain

$$\begin{aligned} D_h F^0(h_\alpha) = D_h\mathcal{L}_0 + \alpha(&D^2_{hx}\mathcal{L}_0 y_{h_0,g} + D^2_{hu}\mathcal{L}_0 v_{h_0,g} + D^2_{hp}\mathcal{L}_0 q_{h_0,g} \\ &+ D^2_{h\varrho}\mathcal{L}_0\varrho_{h_0,g} + D^2_{h\nu}\mathcal{L}_0\mu_{h_0,g} + D^2_{hh}\mathcal{L}_0 g) + o(\alpha\|g\|_H). \end{aligned} \tag{8.6}$$

Substituting (8.6) to (8.5) and integrating, we get

$$\begin{aligned} F^0(h) = F^0(h_0) + D_h\mathcal{L}_0 g + \tfrac{1}{2}(&D^2_{hx}\mathcal{L}_0 y_{h_0,g} + D^2_{hu}\mathcal{L}_0 v_{h_0,g} \\ &+ D^2_{hp}\mathcal{L}_0 q_{h_0,g} + D^2_{h\varrho}\mathcal{L}_0\varrho_{h_0,g} + D^2_{h\nu}\mathcal{L}_0\mu_{h_0,g} + D^2_{hh}\mathcal{L}_0 g, g) + o(\|g\|^2_H). \end{aligned} \tag{8.7}$$

Stationarity conditions for $(\mathrm{L})_{h_0,g}$ have the form

$$\begin{cases} D^2_{xx}\mathcal{L}_0 y_{h_0,g} + D^2_{xu}\mathcal{L}_0 v_{h_0,g} + D^2_{xp}\mathcal{L}_0 q_{h_0,g} + D^2_{x\varrho}\mathcal{L}_0\varrho_{h_0,g} \\ \qquad + D^2_{x\nu}\mathcal{L}_0\mu_{h_0,g} + D^2_{xh}\mathcal{L}_0 g = 0, \\ D^2_{ux}\mathcal{L}_0 y_{h_0,g} + D^2_{uu}\mathcal{L}_0 v_{h_0,g} + D^2_{up}\mathcal{L}_0 q_{h_0,g} + D^2_{u\nu}\mathcal{L}_0\mu_{h_0,g} + D^2_{uh}\mathcal{L}_0 g = 0. \end{cases} \tag{8.8}$$

Moreover, equality constraints and complementarity condition for $(\mathrm{L})_{h_0,g}$ yield

$$\begin{cases} D^2_{px}\mathcal{L}_0 y_{h_0,g} + D^2_{pu}\mathcal{L}_0 v_{h_0,g} + D^2_{ph}\mathcal{L}_0 g = 0, \\ D^2_{\varrho x}\mathcal{L}_0 y_{h_0,g} + D^2_{\varrho h}\mathcal{L}_0 g = 0, \\ (\mu_{h_0,g}, D^2_{\nu x}\mathcal{L}_0 y_{h_0,g} + D^2_{\nu u}\mathcal{L}_0 v_{h_0,g} + D^2_{\nu h}\mathcal{L}_0 g) = 0. \end{cases} \tag{8.9}$$

A combination of (8.7)–(8.9) gives (8.3). ∎

Bibliographical note. This section is based on [37]. Fréchet differentiability of solutions to $(\mathrm{O})_h$ was investigated in [39] using the so-called *shooting method*. Since this method is based on the classical implicit function theorem, it requires strong regularity assumptions on the reference solution, which are not needed in Corollary 8.2.

9. Necessary conditions of stability and sensitivity

In this section we show that if the standing assumptions (I)–(III) hold, and if in addition the dependence of the data of problem $(\mathrm{O})_h$ on the parameter h is *strong* in some sense, then (A1)–(A3) are *necessary* conditions of local Lipschitz continuity and directional differentiability of the solutions and Lagrange multipliers. Thus, in this case, we get a full characterization of stability and sensitivity properties of $(\mathrm{O})_h$.

Let us formulate the needed condition of *strong dependence* of data on the parameter for the abstract generalized equation.

Definition 9.1. We say that the parametric generalized equation (4.5) depends *strongly* on the parameter h at the reference point (h_0, ζ_0) if the linear mapping

$$D_h\mathcal{F}(\zeta_0, h_0) : H \to U^\infty \tag{9.1}$$

is surjective.

Remark 9.2. The strong dependence condition is clearly satisfied in the special situation where

$$\begin{cases} H = H' \times H'', & \text{with } H'' = U^\infty,\ H' \text{ an arbitrary Banach space,} \\ h = (h', h''), & \text{with } h' \in H',\ h'' \in H'', \\ \mathcal{F}(\zeta, h) = \mathcal{F}'(\zeta, h') + h'', \end{cases} \tag{9.2}$$

where $\mathcal{F}'$ satisfies the same assumptions as $\mathcal{F}$ in Theorem 4.1. This special situation was considered in [16].

In the case where $\mathcal{F}$ is given by (4.6), we have

$$D_h\mathcal{F}(\zeta_0, h_0) = \begin{pmatrix} -D_h f(x_0, u_0, h_0) \\ D_h\xi(x_0(0), x_0(1), h_0(0)) \\ D_h\theta(x_0, u_0, h_0) \\ D^2_{xh}\mathcal{H}(x_0, u_0, p_0, \nu_0, h_0) \\ D^2_{x(0)h}[\xi(x_0(0), x_0(1), h_0(0))^*\varrho_0 + \psi(x_0(0), x_0(1), h_0(0)] \\ D^2_{x(1)h}[\xi(x_0(0), x_0(1), h_0(0))^*\varrho_0 + \psi(x_0(0), x_0(1), h_0(0)] \\ D^2_{uh}\mathcal{H}(x_0, u_0, p_0, \nu_0, h_0) \end{pmatrix}. \tag{9.3}$$

We assume

(A4) The mapping $D_h\mathcal{F}(\zeta_0, h_0) : H \to U^\infty$ given in (9.3) is surjective, i.e., by the Banach open mapping theorem, there exists a constant $k > 0$ such that for each $\delta \in U^\infty$ there is $g \in H$ such that

$$D_h\mathcal{F}(\zeta_0, h_0)g = \delta \quad \text{and} \quad \|g\|_H \le k\|\delta\|_{U^\infty}. \tag{9.4}$$

Note that the special case (9.2), where (A4) holds automatically, corresponds to the situation where $H = H' \times H''$ with

$$H'' = L^\infty(0,1;\mathbb{R}^n) \times \mathbb{R}^n \times L^\infty(0,1;\mathbb{R}^l) \times L^\infty(0,1;\mathbb{R}^n) \times \mathbb{R}^n \times \mathbb{R}^n \times L^\infty(0,1;\mathbb{R}^m),$$
$$h = (h', h'') \quad \text{with} \quad h'' = (h''_1, h''_2, h''_3, h''_4, h''_5, h''_6, h''_7) \in H''.$$

In this case problem $(\mathrm{O})_h$ takes the form: Minimize

$$F(x,u,h)=\int_0^1[\varphi(x,u,h')+\langle h_4'',x\rangle+\langle h_7'',u\rangle]\,dt$$
$$+\psi(x(0),x(1),h'(0))+\langle h_5'',x(0)\rangle+\langle h_6'',x(1)\rangle$$

subject to

$$\dot{x}-f(x,u,h')+h_1''=0,$$
$$\xi(x(0),x(1),h'(0))+h_2''=0,$$
$$\theta(x,u,h')+h_3''\le 0.$$

In what follows we assume that

(H) Conditions (I)–(III), (A4) hold and there exist constants $\sigma_1,\sigma_2,l>0$ such that for any $h\in\mathcal{B}^H_{\sigma_1}(h_0)$ there is a unique solution ζ_h in $\mathcal{B}^{W^\infty}_{\sigma_2}(\zeta_0)$ to (4.5) which corresponds to a solution and Lagrange multiplier of $(\mathrm{O})_h$. Moreover, ζ_h is a Lipschitz continuous (with modulus l) and B-differentiable function of h.

We will show that if (H) is satisfied then (A1)–(A3) hold with some $\alpha>0$. To this end, we introduce some variations of the reference value h_0 of the parameter. Namely, we consider the equation

$$\mathcal{F}(\zeta_0,h)=\mathcal{F}(\zeta_0,h_0)+\delta, \tag{9.5}$$

where h,ζ_0 and δ are treated as the unknown, parameter and perturbation, respectively. Certainly, for $\delta=0$, h_0 satisfies (9.5). Hence, in view of (A4), the Lyusternik–Graves theorem 4.3 implies that, for any δ sufficiently small, there exists a solution h_δ of (9.5) such that

$$\|h_\delta-h_0\|_H\le k\|\delta\|_{U^\infty}, \tag{9.6}$$

where $k>0$ is given in (9.4). Therefore, for $\|\delta\|_{U^\infty}<k^{-1}\sigma_1$, we have $h_\delta\in\mathcal{B}^H_{\sigma_1}(h_0)$ and, by (H), there exists a locally unique solution ζ_{h_δ} of the generalized equation

$$0\in\mathcal{F}(\zeta,h_\delta)+\mathcal{T}(\zeta). \tag{9.7}$$

In the proof of constraint qualifications, we choose a small perturbation $\widehat{\delta}$ and a corresponding parameter $\widehat{h}:=h_{\widehat{\delta}}$ in such a way that, in a small neighborhood of $(\widehat{h},\zeta_{\widehat{h}})$, the inclusion (4.5) reduces to an equation. Analyzing this equation, we prove that constraint qualifications (A1) and (A2) hold at $(\widehat{h},\zeta_{\widehat{h}})$. Making $\widehat{h}$ sufficiently close to h_0, we show that these constraint qualifications are satisfied also at the reference solution. Using a similar approach, with a different choice of the perturbation, we prove coercivity condition (A3).

We will need the following well known stability result for surjectivity. A proof is given for the sake of completeness.

LEMMA 9.3. *If* (H) *holds, then there exist* $\sigma_1,\sigma_2>0$ *such that for any* $h\in\mathcal{B}^H_{\sigma_1}(h_0)$, $\zeta\in\mathcal{B}^{W^\infty}_{\sigma_2}(\zeta_0)$ *and any* $\delta\in U^\infty$ *there is* $g\in H$ *such that*

$$D_h\mathcal{F}(\zeta,h)g=\delta\quad\text{and}\quad\|g\|_H\le 2k\|\delta\|_{U^\infty}, \tag{9.8}$$

where k *is given in* (9.4).

Proof. Choose $\sigma_1, \sigma_2 > 0$ so small that

$$\|D_h\mathcal{F}(\zeta,h) - D_h\mathcal{F}(\zeta_0,h_0)\| \le (2k)^{-1} \quad \text{for all } h \in \mathcal{B}^H_{\sigma_1}(h_0) \text{ and } \zeta \in \mathcal{B}^{W^\infty}_{\sigma_2}(\zeta_0). \tag{9.9}$$

Using (9.4), for any $\delta \in U^\infty$, we can construct successively a sequence $\{\sum_{s=0}^j g_s\}$, where

$$\begin{cases} D_h\mathcal{F}(\zeta_0,h_0)g_0 = \delta, \quad \|g_0\|_H \le k\|\delta\|_{U^\infty}, \\ D_h\mathcal{F}(\zeta_0,h_0)g_{j+1} = [D_h\mathcal{F}(\zeta_0,h_0) - D_h\mathcal{F}(\zeta,h)]g_j, \\ \|g_{j+1}\|_H \le k\|[D_h\mathcal{F}(\zeta_0,h_0) - D_h\mathcal{F}(\zeta,h)]g_j\|_{U^\infty}. \end{cases} \tag{9.10}$$

From (9.10) and (9.9) we get

$$\|g_{j+1}\|_H \le k\|D_h\mathcal{F}(\zeta_0,h_0) - D_h\mathcal{F}(\zeta,h)\| \cdot \|g_j\|_H \le \tfrac{1}{2}\|g_j\|_H.$$

So

$$\|g_j\|_H \to 0 \quad \text{as } j \to \infty \tag{9.11}$$

and $\{\sum_{s=0}^j g_s\}$ is a Cauchy sequence. Hence there exists $\overline{g} \in H$ such that

$$\sum_{s=0}^{j} g_s \to \overline{g} \quad \text{as } j \to \infty \quad \text{and} \quad \|\overline{g}\|_H \le 2k\|\delta\|_{U^\infty}. \tag{9.12}$$

On the other hand, it follows by induction that

$$D_h\mathcal{F}(\zeta,h)\Big(\sum_{s=0}^{j} g_s\Big) - \delta = -D_h\mathcal{F}(\zeta_0,h_0)g_{j+1}.$$

In view of (9.11) and (9.12), this shows that $\overline{g}$ satisfies (9.8). ■

Let us now analyze constraint qualifications. To this end, we choose small constants $\alpha, \varepsilon > 0$ and introduce the following variation $\widehat{\nu}$ of ν_0:

$$\widehat{\nu}^i(t) = \begin{cases} \nu_0^i(t) + \varepsilon & \text{if } t \in M_\alpha^i, \\ \nu_0^i(t) = 0 & \text{if } t \notin M_\alpha^i, \end{cases} \tag{9.13}$$

where M_α^i is defined in (2.38). Set $\widehat{\zeta} = (x_0, u_0, p_0, \varrho_0, \widehat{\nu})$. Moreover, define the following vector $\widehat{\delta} \in U^\infty$:

$$\begin{cases} \widehat{\delta} = (0,0,\widehat{\delta}^3,\widehat{\delta}^4,0,0,\widehat{\delta}^7), \quad \text{where} \\ (\widehat{\delta}^3)^i(t) = \begin{cases} -\theta_0^i(t) & \text{if } i \in I_\alpha(t), \\ 0 & \text{if } i \notin I_\alpha(t), \end{cases} \\ (\widehat{\delta}^4)^i(t) = -\varepsilon \sum_{j\in I_\alpha(t)} \Upsilon^{ji}(t), \\ (\widehat{\delta}^7)^i(t) = -\varepsilon \sum_{j\in I_\alpha(t)} \Theta^{ji}(t). \end{cases} \tag{9.14}$$

It follows from the construction that

$$0 \in \mathcal{F}(\widehat{\zeta},h_0) + \widehat{\delta} + \mathcal{T}(\widehat{\zeta}) \tag{9.15}$$

and (see Fig. 1)

$$\begin{cases} \theta^i(x_0(t),u_0(t),h_0(t)) + (\widehat{\delta}^3)^i(t) = 0 & \text{and} \quad \widehat{\nu}^i(t) \ge \varepsilon \text{ for } t \in M_\alpha^i, \\ \theta^i(x_0(t),u_0(t),h_0(t)) + (\widehat{\delta}^3)^i(t) < -\alpha & \text{and} \quad \widehat{\nu}^i(t) = 0 \text{ for } t \notin M_\alpha^i. \end{cases} \tag{9.16}$$

Clearly, by (9.13) and (9.14), we have

$$\|\widehat{\zeta} - \zeta_0\|_{W^\infty} \le \varepsilon, \quad \|\widehat{\delta}^3\|_\infty \le \alpha, \quad \|\widehat{\delta}^4\|_\infty \le \varepsilon\|\Upsilon\|_\infty, \quad \|\widehat{\delta}^7\|_\infty \le \varepsilon\|\Theta\|_\infty. \tag{9.17}$$

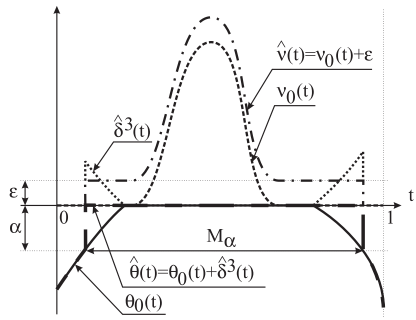

Fig. 1

So, in view of Lemma 9.3 and Theorem 4.3, for $\alpha > 0$ and $\varepsilon > 0$ sufficiently small the equation

$$\mathcal{F}(\widehat{\zeta}, h) = \mathcal{F}(\widehat{\zeta}, h_0) + \widehat{\delta} \tag{9.18}$$

has a solution $\widehat{h}$ such that

$$\|\widehat{h} - h_0\| \leq 2k\|\widehat{\delta}\|_{U^\infty}, \tag{9.19}$$

where $k > 0$ is independent of α and ε. By (9.15) and (9.18) we get

$$0 \in \mathcal{F}(\widehat{\zeta}, \widehat{h}) + \mathcal{T}(\widehat{\zeta}). \tag{9.20}$$

Therefore, by (H), $\widehat{\zeta} = \zeta_{\widehat{h}}$ is a unique solution in $\mathcal{B}_{\sigma_2}^{Z^\infty}(\zeta_0)$ of (9.20). By (9.16) and (9.18) we have

$$\begin{cases} \theta^i(x_0(t), u_0(t), \widehat{h}(t)) = 0 & \text{and} \quad \widehat{\nu}^i(t) \geq \varepsilon \text{ for } t \in M_\alpha^i, \\ \theta^i(x_0(t), u_0(t), \widehat{h}(t)) < -\alpha & \text{and} \quad \widehat{\nu}^i(t) = 0 \text{ for } t \notin M_\alpha^i. \end{cases} \tag{9.21}$$

Conditions (9.21), together with the assumption of Lipschitz continuity of ζ_h, imply that there exists a small constant $\widehat{\sigma} > 0$ such that, for all $h \in \mathcal{B}_{\widehat{\sigma}}^H(\widehat{h}) \subset \mathcal{B}_{\sigma_1}^H(h_0)$ and all $\zeta \in \mathcal{B}_{\sigma_2}^{Z^\infty}(\zeta_0)$, the generalized equation (4.5) can be treated as the following *equation*:

$$\widehat{\mathcal{F}}(\zeta, h) = 0, \tag{9.22}$$

where $\widehat{\mathcal{F}} : X^s \to \widehat{U}_\alpha^s$ is defined as in (4.6) except for the third term:

$$\widehat{\mathcal{F}}(\zeta, h) = \begin{pmatrix} \dot{x} - f(x, u, h) \\ \xi(x(0), x(1), h(0)) \\ \theta^i(x(t), u(t), h(t)) \quad \text{for } t \in M_\alpha^i, \ i \in I \\ \dot{p} + D_x\mathcal{H}(x, u, p, \nu(t), h) \\ p(0) + D_{x(0)}[\xi(x(0), x(1), h(0))^*\varrho + \psi(x(0), x(1), h(0))] \\ -p(1) + D_{x(1)}[\xi(x(0), x(1), h(0))^*\varrho + \psi(x(0), x(1), h(0))] \\ D_u\mathcal{H}(x, u, p, \nu, h) \end{pmatrix}. \tag{9.23}$$

By (H), the solution ζ_h of (9.22) is B-differentiable on $\mathcal{B}_{\widehat{\sigma}}^H(\widehat{h})$. Let $g \in H$ be any direction in the space of parameters. Differentiating (9.22) at $h = \widehat{h}$ along the direction g we obtain

$$D_\zeta\widehat{\mathcal{F}}(\zeta_{\widehat{h}}, \widehat{h})(D_h\zeta_{\widehat{h}}, g) = -D_h\widehat{\mathcal{F}}(\zeta_{\widehat{h}}, \widehat{h})g. \tag{9.24}$$

For any $g \in H$ this equation must have a unique solution $(D_h\zeta_{\widehat{h}}, g)$, and, by Lipschitz continuity,

$$\|(D_h\zeta_{\widehat{h}}, g)\|_{W^\infty} \leq l\|g\|_H. \tag{9.25}$$

By Lemma 9.3, for $\alpha > 0$ and $\varepsilon > 0$ sufficiently small, $D_h\mathcal{F}(\zeta_{\widehat{h}}, \widehat{h}) : H \to U^\infty$ is surjective and condition (9.8) is satisfied. Hence, it follows from definition (9.23) that the mapping $D_h\widehat{\mathcal{F}}(\zeta_{\widehat{h}}, \widehat{h}) : H \to \widehat{U}^\infty_\alpha$ is surjective. Therefore, the range of the right-hand side of (9.24) is the whole space $\widehat{U}^\infty_\alpha$. This implies that $D_\zeta\widehat{\mathcal{F}}(\zeta_{\widehat{h}}, \widehat{h}) : Z^\infty \to \widehat{U}^\infty_\alpha$ is invertible and

$$(D_h\zeta_{\widehat{h}}, g) = -[D_\zeta\widehat{\mathcal{F}}(\zeta_{\widehat{h}}, \widehat{h})]^{-1} D_h\widehat{\mathcal{F}}(\zeta_{\widehat{h}}, \widehat{h})g. \tag{9.26}$$

For any $\varepsilon > 0$, we can choose an element $\delta_\varepsilon \in \widehat{U}^\infty_\alpha$, $\|\delta_\varepsilon\|_{\widehat{U}^\infty_\alpha} = 1$, such that

$$\|[D_\zeta\widehat{\mathcal{F}}(\zeta_{\widehat{h}}, \widehat{h})]^{-1}\delta_\varepsilon\|_{W^\infty} \geq \|[D_\zeta\widehat{\mathcal{F}}(\zeta_{\widehat{h}}, \widehat{h})]^{-1}\|_{\widehat{U}^\infty_\alpha \to W^\infty}\|\delta_\varepsilon\|_{\widehat{U}^\infty_\alpha} - \varepsilon. \tag{9.27}$$

In view of (9.8), there exists $g_\varepsilon \in H$ such that

$$D_h\widehat{\mathcal{F}}(\zeta_{\widehat{h}}, \widehat{h})g_\varepsilon = \delta_\varepsilon \quad \text{and} \quad \|g_\varepsilon\|_H \leq 2k\|\delta_\varepsilon\|_{\widehat{U}^\infty_\alpha}. \tag{9.28}$$

By (9.26)–(9.28),

$$\begin{aligned} \|(D_h\zeta_{\widehat{h}}, g_\varepsilon)\|_{W^\infty} &= \|[D_\zeta\widehat{\mathcal{F}}(\zeta_{\widehat{h}}, \widehat{h})]^{-1} D_h\widehat{\mathcal{F}}(\zeta_{\widehat{h}}, \widehat{h})g_\varepsilon\|_{W^\infty} \\ &\geq \frac{1}{2k}\|[D_\zeta\widehat{\mathcal{F}}(\zeta_{\widehat{h}}, \widehat{h})]^{-1}\|_{\widehat{U}^\infty_\alpha \to W^\infty}\|g_\varepsilon\|_H - \varepsilon. \end{aligned} \tag{9.29}$$

Since $\varepsilon > 0$ can be arbitrarily small, (9.29) together with (9.25) implies

$$\|[D_\zeta\widehat{\mathcal{F}}(\zeta_{\widehat{h}}, \widehat{h})]^{-1}\|_{\widehat{U}^\infty_\alpha \to W^\infty} \leq 2kl. \tag{9.30}$$

In view of (H), (9.17) and (9.19), for $\alpha > 0$ and $\varepsilon > 0$ sufficiently small, we obtain

$$\|[D_\zeta\widehat{\mathcal{F}}(\zeta_{h_0}, h_0)]^{-1}\|_{\widehat{U}^\infty_\alpha \to W^\infty} \leq 4kl. \tag{9.31}$$

PROPOSITION 9.4. *If* (H) *is satisfied*, *then there exists* $\alpha > 0$ *such that* (A1) *and* (A2) *hold.*

Proof. By (9.23), the equation $D_\zeta\widehat{\mathcal{F}}(\zeta_0, h_0)\eta = \chi$ has the form

$$\begin{cases} \dot{y}(t) - A(t)y(t) - B(t)v(t) = \chi^1(t), \\ \Xi_0 y(0) + \Xi_1 y(1) = \chi^2, \\ \widehat{\Upsilon}_\alpha(t)y(t) + \widehat{\Theta}_\alpha(t)v(t) = \chi^3(t), \\ \dot{q}(t) + A(t)^*q(t) + Q_{11}(t)y(t) + Q_{12}(t)v(t) + \widehat{\Upsilon}_\alpha(t)^*\mu(t) = \chi^4(t), \\ q(0) + \mathcal{R}_{00}y(0) + \mathcal{R}_{01}y(1) + \Xi_0^*\varrho = \chi^5, \\ -q(1) + \mathcal{R}_{10}y(0) + \mathcal{R}_{11}y(1) + \Xi_1^*\varrho = \chi^6, \\ Q_{21}(t)y(t) + Q_{22}(t)v(t) + B(t)^*q(t) + \widehat{\Theta}_\alpha(t)^*\mu(t) = \chi^7(t), \end{cases} \tag{9.32}$$

where $\widehat{\Theta}_\alpha(t)$ and $\widehat{\Upsilon}_\alpha(t)$ are defined in (2.14).

By (9.31), equation (9.32) has a unique solution for any $\chi \in \widehat{U}^\infty_\alpha$. Choosing $\chi = (0, \chi^2, 0, 0, 0, 0, 0)$ we immediately obtain (A2). We are going to show that

$$|\widehat{\Theta}_\alpha(t)^*\mu| \geq (4kl)^{-1}|\mu| \quad \text{for all } \mu \in \mathbb{R}^{\imath_\alpha(t)} \text{ and a.a. } t \in [0,1]. \tag{9.33}$$

In view of Lemma 2.1, (9.33) is equivalent to (A1). Suppose that (9.33) is violated, i.e., there exist a set $M \subset [0,1]$ with $\operatorname{meas} M > 0$, a constant $\varepsilon > 0$ and $\breve{\mu}(t)$, $|\breve{\mu}(t)| = 1$, such that

$$|\widehat{\Theta}_\alpha(t)^*\check{\mu}(t)| \le (4kl)^{-1} - \varepsilon \quad \text{for all } t \in M. \tag{9.34}$$

Let $N \subset M$ with $\operatorname{meas} N > 0$ be any subset. We set

$$\widetilde{\mu}(t) = \begin{cases} \check{\mu}(t) & \text{for } t \in N, \\ 0 & \text{for } t \notin N. \end{cases} \tag{9.35}$$

We can assume that $\widetilde{\mu}$ is a measurable function. Choose

$$\widetilde{\eta} = (\widetilde{y}, \widetilde{u}, \widetilde{q}, \widetilde{\varrho}, \widetilde{\mu}), \quad \widetilde{\chi} = (0, 0, 0, 0, \widetilde{\chi}^5, 0, \widetilde{\chi}^7),$$

where $\widetilde{y} = 0$, $\widetilde{u} = 0$, $\widetilde{\varrho} = 0$, $\widetilde{q}$ is the solution of the equation

$$\dot{\widetilde{q}}(t) + A(t)^*\widetilde{q}(t) + \widehat{\Upsilon}_\alpha(t)^*\widetilde{\mu}(t) = 0, \quad \widetilde{q}(1) = 0, \tag{9.36}$$

and

$$\widetilde{\chi}^5 = \widetilde{q}(0), \quad \widetilde{\chi}^7(t) = B(t)^*\widetilde{q}(t) + \widehat{\Theta}_\alpha(t)^*\widetilde{\mu}(t). \tag{9.37}$$

By construction, $\widetilde{\eta}$ is the solution to (9.32) corresponding to $\widetilde{\chi}$. By (9.35) and (9.36) we have $\|\widetilde{q}\|_\infty \to 0$ as $\operatorname{meas} N \to 0$. Hence, in view of (9.34) and (9.37), for $\operatorname{meas} N$ sufficiently small, we get

$$\begin{aligned} \|\widetilde{\chi}\|_{\widehat{U}^\infty_\alpha} &= \|\widetilde{\chi}^7\|_\infty \le \|\widehat{\Theta}_\alpha(t)^*\widetilde{\mu}(t)\|_\infty + \|B(t)^*\widetilde{q}(t)\|_\infty \\ &\le (4kl)^{-1} - \varepsilon/2 = [(4kl)^{-1} - \varepsilon/2]\|\widetilde{\mu}\|_\infty = [(4kl)^{-1} - \varepsilon/2]\|\widetilde{\eta}\|_{W^\infty}. \end{aligned}$$

This contradicts (9.31) and completes the proof of (9.33). ■

In proving (A3), we will use the same idea as above. Namely, we introduce a variation of h_0 such that, for the corresponding problem, locally, the constraints can be treated as equalities. As in (9.13) we introduce the following variation $\overline{\nu}$ of ν_0:

$$\overline{\nu}^i(t) = \begin{cases} \nu_0^i(t) & \text{if } t \in N^i_\alpha, \\ 0 & \text{if } t \notin N^i_\alpha, \end{cases}$$

where N^i_α is defined in (3.16), and set $\overline{\zeta} = (x_0, u_0, p_0, \varrho_0, \overline{\nu})$. Moreover, as in (9.14), for given $\alpha, \varepsilon > 0$, we define

$$\begin{cases} \overline{\delta} = (0, 0, \overline{\delta}^3, \overline{\delta}^4, 0, 0, \overline{\delta}^7), \quad \text{where} \\ (\overline{\delta}^3)^i(t) = \begin{cases} 0 & \text{if } i \in J_\alpha(t), \\ -\varepsilon & \text{if } i \notin J_\alpha(t), \end{cases} \\ (\overline{\delta}^4)^i(t) = \sum_{j \in J_\alpha(t)} \nu_0^j(t)\Upsilon^{ji}(t), \\ (\overline{\delta}^7)^i(t) = \sum_{j \in J_\alpha(t)} \nu_0^j(t)\Theta^{ji}(t). \end{cases} \tag{9.38}$$

By the same argument as in (9.15)–(9.20), we find that for $\alpha > 0$ and $\varepsilon > 0$ sufficiently small there exists $\overline{h} \in \mathcal{B}^H_{\sigma_1}(h_0)$ such that $\overline{\zeta} = \zeta_{\bar{h}}$ is a locally unique solution to the generalized equation

$$0 \in \mathcal{F}(\zeta, \overline{h}) + \mathcal{T}(\zeta). \tag{9.39}$$

Moreover (see Fig. 2), there exists $\overline{\sigma}$ such that, for all $h \in \mathcal{B}^H_{\overline{\sigma}}(\overline{h})$ and all $\zeta \in \mathcal{B}^{Z^\infty}_{\sigma_2}(\zeta_0)$, (4.5) can be treated as the equation

$$\overline{\mathcal{F}}(\zeta, h) = 0, \tag{9.40}$$

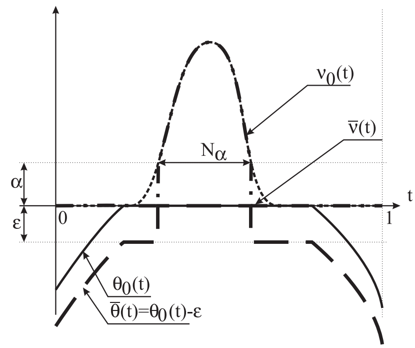

Fig. 2

where $\overline{\mathcal{F}}: X^s \to \overline{U}^s_\alpha$ is defined as in (9.23), with M^i_α replaced by N^i_α. In exactly the same way as in (9.30), we find that

$$\|[D_\zeta \overline{\mathcal{F}}(\zeta_{\bar{h}}, \overline{h})]^{-1}\|_{\overline{U}^\infty_\alpha \to W^\infty} \le 2kl. \tag{9.41}$$

Hence, in particular, for each $\eta \in \overline{U}^\infty_\alpha$, the equation

$$D_\zeta \overline{\mathcal{F}}(\zeta_{\bar{h}}, \overline{h})\zeta = \eta \tag{9.42}$$

has a unique solution and

$$\|\zeta\|_{W^\infty} \le 2kl\|\eta\|_{\overline{U}^\infty_\alpha}. \tag{9.43}$$

Note that (9.42) has the form

$$\begin{cases} \dot{y}(t) - A^{\bar{h}}(t)y(t) - B^{\bar{h}}(t)v(t) = \chi^1(t), \\ \Xi^{\bar{h}}_0 y(0) + \Xi^{\bar{h}}_1 y(1) = \chi^2, \\ \overline{\Upsilon}^{\bar{h}}_\alpha(t)y(t) + \overline{\Theta}^{\bar{h}}_\alpha(t)v(t) = \chi^3(t), \\ \dot{q}(t) + A^{\bar{h}}(t)^* q(t) + Q^{\bar{h}}_{11}(t)y(t) + Q^{\bar{h}}_{12}(t)v(t) + \overline{\Upsilon}^{\bar{h}}_\alpha(t)^*\mu(t) = \chi^4(t), \\ q(0) + \mathcal{R}^{\bar{h}}_{00}y(0) + \mathcal{R}^{\bar{h}}_{01}y(1) + (\Xi^{\bar{h}}_0)^*\varrho = \chi^5, \\ -q(1) + \mathcal{R}^{\bar{h}}_{10}y(0) + \mathcal{R}^{\bar{h}}_{11}y(1) + (\Xi^{\bar{h}}_1)^*\varrho = \chi^6, \\ Q^{\bar{h}}_{21}(t)y(t) + Q^{\bar{h}}_{22}(t)v(t) + B^{\bar{h}}(t)^* q(t) + \overline{\Theta}^{\bar{h}}_\alpha(t)^*\mu(t) = \chi^7(t), \end{cases} \tag{9.44}$$

where all functions are defined as in (2.5), (3.19) and (3.21), but evaluated at $(\zeta_{\bar{h}}, \overline{h})$ rather than at (ζ_0, h_0).

We are going to show that

$$(v, \mathcal{Q}^{\bar{h}}_\alpha v) \ge (2kl)^{-1}\|v\|^2_2 \quad \text{for all } v \in (\mathcal{U}^{\bar{h}}_\alpha)^2, \tag{9.45}$$

where $\mathcal{Q}^h_\alpha$ is defined in (7.9). To this end, note that $\mathcal{Q}^{\bar{h}}_\alpha$ is a self-adjoint operator in $(\mathcal{U}^{\bar{h}}_\alpha)^2$. By a well known property of the spectrum of self-adjoint operators in a Hilbert space (see e.g. Theorem 2, p. 320 in [53]) we have

$$\min\{\lambda \in \mathbb{R} \mid \lambda \in \sigma\} = \inf\{(v, \mathcal{Q}^{\bar{h}}_\alpha v) \mid v \in (\mathcal{U}^{\bar{h}}_\alpha)^2,\ \|v\|_2 = 1\}, \tag{9.46}$$

where σ is the spectrum of $\mathcal{Q}^{\bar{h}}_\alpha$. Hence, to prove (9.45), it is enough to show that

$$\sigma \subset [(2kl)^{-1}, \infty]. \tag{9.47}$$

Let us start with the following result:

LEMMA 9.5. *If* (H) *is satisfied then there is* $\alpha > 0$ *such that*

$$(v, \mathcal{Q}_\alpha^{\bar{h}})v) \geq 0 \quad \textit{for all } v \in (\mathcal{U}_\alpha^{\bar{h}})^2. \tag{9.48}$$

Proof. Locally, $(\mathrm{O})_{\bar{h}}$ can be treated as the following optimal control problem with equality type constraints:

$$\text{Minimize } F(x, u, \overline{h}) \text{ subject to } \overline{c}(x, u, \overline{h}) = 0,$$

where

$$\overline{c}(x, u, h) = \begin{pmatrix} \dot{x} - f(x, u, h) \\ \xi(x(0), x(1), h(0)) \\ \theta^i(x(t), u(t), h(t)) \quad \text{for } t \in N_\alpha^i,\ i \in I \end{pmatrix}.$$

Choose an arbitrary $(y, v) \in X^\infty \cap \ker \overline{\mathcal{C}}_\alpha^{\bar{h}}$, where $\overline{\mathcal{C}}_\alpha^h$ is defined in (7.8). Set $(x^\tau, u^\tau) = (x_0, u_0) + \tau(y, v)$, where $\tau > 0$. Expanding $\overline{c}(\cdot, \cdot, \overline{h})$ into Taylor series at (x_0, u_0) we get

$$\overline{c}(x^\tau, u^\tau, \overline{h}) = \overline{c}(x_0, u_0, \overline{h}) + \tau \overline{\mathcal{C}}_\alpha^{\bar{h}} \begin{bmatrix} y \\ v \end{bmatrix} + o(\tau) = o(\tau). \tag{9.49}$$

In view of Corollary 2.4, the mapping $\widehat{\mathcal{C}}_\alpha$ is surjective for any sufficiently small $\alpha > 0$. Therefore, it follows from the construction of $\overline{\mathcal{C}}_\alpha^{\bar{h}}$ and from the same argument as in the proof of Lemma 9.3 that $\overline{\mathcal{C}}_\alpha^{\bar{h}}$ is surjective for $\alpha > 0$ and $\varepsilon > 0$ sufficiently small. Hence, by the Lyusternik–Graves theorem, (9.49) implies that there exists $(\widetilde{x}^\tau, \widetilde{u}^\tau)$ feasible for $(\mathrm{O})_{\bar{h}}$ such that

$$\|(\widetilde{x}^\tau, \widetilde{u}^\tau) - (x^\tau, u^\tau)\|_{X^\infty} = \|(\widetilde{x}^\tau - x_0 - \tau y, \widetilde{u}^\tau - u_0 - \tau v)\|_{X^\infty} = o(\tau).$$

For any feasible (x, u) we have $F(x, u, \overline{h}) = \mathcal{L}(x, u, p_{\bar{h}}, \varrho_{\bar{h}}, \nu_{\bar{h}}, \overline{h})$. Since $(x_{\bar{h}}, u_{\bar{h}})$ is a local minimizer of $(\mathrm{O})_{\bar{h}}$, expanding $F(\cdot, \cdot, \overline{h})$ into Taylor series at $(x_{\bar{h}}, u_{\bar{h}})$ and performing calculations analogous to those in the proof of Lemma 7.4, we obtain

$$0 \leq F(\widetilde{x}^\tau, \widetilde{u}^\tau, \overline{h}) - F(x_{\bar{h}}, u_{\bar{h}}, \overline{h}) = \tfrac{1}{2}\tau^2((y, v), D^2\mathcal{L}_{\bar{h}}(y, v)) + o(\tau^2).$$

Letting $\tau \to 0$ and using density of the embedding $X^\infty \subset X^2$, we get

$$((y, v), D^2\mathcal{L}_{\bar{h}}(y, v)) \geq 0 \quad \text{for all } (y, v) \in \ker \overline{\mathcal{C}}_\alpha^{\bar{h}}. \tag{9.50}$$

In the same way as in (3.28), we can rewrite (9.50) in the form (9.48). ■

We will need the following auxiliary lemma:

LEMMA 9.6. *If* (H) *is satisfied then*

$$\left| \begin{bmatrix} \overline{\Theta}_\alpha^{\bar{h}}(t) & 0 \\ Q_{22}^{\bar{h}}(t) & \overline{\Theta}_\alpha^{\bar{h}}(t)^* \end{bmatrix} \begin{bmatrix} v \\ \mu \end{bmatrix} \right| \geq (2kl)^{-1} \left| \begin{matrix} v \\ \mu \end{matrix} \right| \tag{9.51}$$

for all (v, μ) *of the appropriate dimension and a.a.* $t \in [0, 1]$.

Proof. The idea of the proof is very similar to that of (9.33). Suppose that (9.51) is violated, i.e., there exist a set $M \subset [0, 1]$ with meas $M > 0$, a constant $\varepsilon > 0$ and vectors $(\check{v}(t), \check{\mu}(t))$, $|(\check{v}(t), \check{\mu}(t))| = 1$, such that

$$\left| \begin{bmatrix} \overline{\Theta}_\alpha^{\bar{h}}(t) & 0 \\ Q_{22}^{\bar{h}}(t) & \overline{\Theta}_\alpha^{\bar{h}}(t)^* \end{bmatrix} \begin{bmatrix} \check{v} \\ \check{\mu} \end{bmatrix} \right| \leq [(2kl)^{-1} - \varepsilon] \left| \begin{matrix} \check{v} \\ \check{\mu} \end{matrix} \right| \quad \text{for all } t \in M. \tag{9.52}$$

Choose any $N \subset M$ with meas $N > 0$ and set

$$(\widetilde{v}(t), \widetilde{\mu}(t)) = \begin{cases} (\check{v}(t), \check{\mu}(t)) & \text{for } t \in N, \\ (0,0) & \text{for } t \notin N. \end{cases} \tag{9.53}$$

We can assume that $(\widetilde{v}, \widetilde{\mu})$ is a measurable function. Define

$$\widetilde{\eta} = (\widetilde{y}, \widetilde{v}, \widetilde{q}, \widetilde{\varrho}, \widetilde{\mu}), \quad \widetilde{\chi} = (0, \widetilde{\chi}^2, \widetilde{\chi}^3, 0, \widetilde{\chi}^5, \widetilde{\chi}^6, \widetilde{\chi}^7),$$

where

$$\begin{cases} \dot{\widetilde{y}} - A^{\bar{h}}\widetilde{y} - B^{\bar{h}}\widetilde{v} = 0, \quad \widetilde{y}(0) = 0, \\ \dot{\widetilde{q}} + (A^{\bar{h}})^*\widetilde{q} + Q^{\bar{h}}_{11}\widetilde{y} + Q^{\bar{h}}_{12}\widetilde{v} + (\overline{\Upsilon}^{\bar{h}})^*\widetilde{\mu} = 0, \quad \widetilde{q}(1) = 0, \\ \widetilde{\varrho} = 0, \\ \widetilde{\chi}^2 = \Xi^{\bar{h}}_1\widetilde{y}(1), \quad \widetilde{\chi}^3 = \overline{\Upsilon}^{\bar{h}}_\alpha\widetilde{y}, + \overline{\Theta}^{\bar{h}}_\alpha\widetilde{v}, \quad \widetilde{\chi}^5 = \widetilde{q}(0) + \mathcal{R}^{\bar{h}}_{01}\widetilde{y}(1), \\ \widetilde{\chi}^6 = \mathcal{R}^{\bar{h}}_{11}\widetilde{y}(1), \quad \widetilde{\chi}^7 = Q^{\bar{h}}_{21}\widetilde{y} + Q^{\bar{h}}_{22}\widetilde{v} + (B^{\bar{h}})^*\widetilde{q} + (\overline{\Theta}^{\bar{h}}_\alpha)^*\widetilde{\mu}. \end{cases} \tag{9.54}$$

It is easy to see that $\widetilde{\eta}$ is the solution of (9.44) corresponding to $\widetilde{\chi}$. It follows from (9.53) and (9.54) that

$$\|\widetilde{y}\|_\infty, \|\widetilde{q}\|_\infty \to 0 \quad \text{as meas } N \to 0.$$

Hence, in view of (9.52) and (9.54), for meas N sufficiently small, we obtain

$$\begin{aligned} \|\widetilde{\chi}\|_{\overline{U}^\infty_\alpha} &= \left\| \begin{bmatrix} \widetilde{\chi}^3 \\ \widetilde{\chi}^7 \end{bmatrix} \right\|_\infty \le \left\| \begin{bmatrix} \overline{\Theta}^{\bar{h}}_\alpha & 0 \\ Q^{\bar{h}}_{22} & (\overline{\Theta}^{\bar{h}}_\alpha)^* \end{bmatrix} \begin{bmatrix} \widetilde{v} \\ \widetilde{\mu} \end{bmatrix} \right\|_\infty + \left\| \begin{matrix} \overline{\Upsilon}^{\bar{h}}_\alpha \widetilde{y} \\ Q^{\bar{h}}_{12}\widetilde{y} + (B^{\bar{h}})^*\widetilde{q} \end{matrix} \right\|_\infty \\ &\le [(2kl)^{-1} - \varepsilon] \left\| \begin{matrix} \widetilde{v} \\ \widetilde{\mu} \end{matrix} \right\|_\infty + \varepsilon/2 \le (2kl)^{-1} - \varepsilon/2 = [(2kl)^{-1} - \varepsilon/2]\|\widetilde{\eta}\|_{W^\infty}. \end{aligned}$$

This contradicts (9.43) and completes the proof of the lemma. ∎

LEMMA 9.7. *If* (H) *is satisfied*, *then there exists* $\alpha > 0$ *such that*

$$\|\mathcal{Q}^{\bar{h}}_\alpha v\|_2 \ge (2kl)^{-1}\|v\|_2 \quad \textit{for all } v \in (\mathcal{U}^{\bar{h}}_\alpha)^2. \tag{9.55}$$

Proof. Set $\chi = (0, 0, \chi^3, 0, 0, 0, \chi^7)$. Using definitions (3.23) and (3.25) we obtain from (9.44),

$$\begin{aligned} \overline{\Theta}^{\bar{h}}_\alpha v + \overline{\Upsilon}^{\bar{h}}\mathcal{S}^{\bar{h}}_0 B^{\bar{h}} v &= \chi^3, \\ Q^{\bar{h}}_{21}\mathcal{S}^{\bar{h}}_0 B^{\bar{h}} v + (B^{\bar{h}})^*(\mathcal{S}^{\bar{h}})^* \begin{pmatrix} Q^{\bar{h}}_{11} & 0 & 0 \\ 0 & \mathcal{R}^{\bar{h}}_{00} & \mathcal{R}^{\bar{h}}_{01} \\ 0 & \mathcal{R}^{\bar{h}}_{10} & \mathcal{R}^{\bar{h}}_{11} \end{pmatrix} &\mathcal{S}^{\bar{h}} B^{\bar{h}} v + (B^{\bar{h}})^*(\mathcal{S}^{\bar{h}}_0)^* Q^{\bar{h}}_{12} v \\ &+ Q^{\bar{h}}_{22} v + (B^{\bar{h}})^*(\mathcal{S}^{\bar{h}}_0)^*(\overline{\Upsilon}^{\bar{h}}_\alpha)^*\mu + (\overline{\Theta}^{\bar{h}}_\alpha)^*\mu \\ &= Q^{\bar{h}}_{22} v + (\overline{\Theta}^{\bar{h}}_\alpha)^*\mu + \mathcal{M}^{\bar{h}} v + (B^{\bar{h}})^*(\mathcal{S}^{\bar{h}}_0)^*(\overline{\Upsilon}^{\bar{h}}_\alpha)^*\mu = \chi^7, \end{aligned}$$

or

$$\begin{bmatrix} \overline{\Theta}^{\bar{h}}_\alpha & 0 \\ Q^{\bar{h}}_{22} & (\overline{\Theta}^{\bar{h}}_\alpha)^* \end{bmatrix} \begin{bmatrix} v \\ \mu \end{bmatrix} + \begin{bmatrix} \overline{\Upsilon}^{\bar{h}}_\alpha\mathcal{S}^{\bar{h}}_0 B^{\bar{h}} & 0 \\ \mathcal{M}^{\bar{h}} & (B^{\bar{h}})^*(\mathcal{S}^{\bar{h}}_0)^*(\overline{\Upsilon}^{\bar{h}}_\alpha)^* \end{bmatrix} \begin{bmatrix} v \\ \mu \end{bmatrix} = \begin{bmatrix} \chi^3 \\ \chi^7 \end{bmatrix}. \tag{9.56}$$

Introduce the space $V^s_\alpha = \overline{L}^s_\alpha \times L^s(0,1;\mathbb{R}^m)$, where $\overline{L}^s_\alpha$ is defined in (3.16). By (9.43), for any $(\chi^3, \chi^7) \in V^\infty_\alpha$, equation (9.56) has a unique solution and

$$\left\| \begin{matrix} v \\ \mu \end{matrix} \right\|_\infty \le 2kl \left\| \begin{matrix} \chi^3 \\ \chi^7 \end{matrix} \right\|_\infty. \tag{9.57}$$

We are going to show that the analogous estimate holds in V^2_α. Define the operator

$$\mathcal{D} : V^s_\alpha \to V^s_\alpha, \quad \mathcal{D} = \mathcal{D}_1 + \mathcal{D}_2, \tag{9.58}$$

where

$$\mathcal{D}_1 = \begin{bmatrix} \overline{\Theta}^{\bar h}_\alpha & 0 \\ Q^{\bar h}_{22} & (\overline{\Theta}^{\bar h}_\alpha)^* \end{bmatrix}, \quad \mathcal{D}_2 = \begin{bmatrix} \overline{\Upsilon}^{\bar h} \mathcal{S}^{\bar h}_0 B^{\bar h} & 0 \\ \mathcal{M}^{\bar h} & (B^{\bar h})^* (\mathcal{S}^{\bar h}_0)^* (\overline{\Upsilon}^{\bar h})^* \end{bmatrix}.$$

Let $\mathcal{I}$ denote the identity in V^s_α. In view of (9.57), for any $\lambda \in (-(2kl)^{-1}, (2kl)^{-1})$, the operator

$$\mathcal{D} - \lambda \mathcal{I} : V^s_\alpha \to V^s_\alpha \tag{9.59}$$

is invertible for $s = \infty$. We are going to prove that (9.59) is also invertible for $s = 2$. To this end, notice that by (9.51), for any $\lambda \in (-(2kl)^{-1}, (2kl)^{-1})$, the matrix

$$\begin{bmatrix} \overline{\Theta}^{\bar h}_\alpha(t) & 0 \\ Q^{\bar h}_{22}(t) & \overline{\Theta}^{\bar h}_\alpha(t)^* \end{bmatrix} - \lambda I$$

is invertible, uniformly for almost all $t \in [0,1]$. Therefore, the operator $\mathcal{D}_1 - \lambda\mathcal{I}$ is invertible in V^s_α for any $s \in [1, \infty]$. Define the operator

$$\mathcal{E}_\lambda : V^s_\alpha \to V^s_\alpha, \quad \mathcal{E}_\lambda = (\mathcal{D}_1 - \lambda\mathcal{I})^{-1}\mathcal{D}_2.$$

Observe that $\mathcal{D} - \lambda\mathcal{I}$ is invertible in V^s_α if and only if $\mathcal{E}_\lambda + \mathcal{I}$ is also invertible. In particular, $\mathcal{E}_\lambda + \mathcal{I}$ is invertible in V^∞_α for $\lambda \in (-(2kl)^{-1}, (2kl)^{-1})$. Note that, in view of (3.23) and (3.25), the mappings $\mathcal{S}^{\bar h}_0 : L^2(0,1;\mathbb{R}^m) \to L^2(0,1;\mathbb{R}^n)$ and $\mathcal{M}^{\bar h} : L^2(0,1;\mathbb{R}^m) \to L^2(0,1;\mathbb{R}^m)$ are compact. Hence $\mathcal{D}_2$ and $\mathcal{E}_\lambda$ are compact in V^2_α. Moreover,

$$\mathcal{E}_\lambda V^2_\alpha \subset V^\infty_\alpha. \tag{9.60}$$

Consider the homogeneous equation

$$(\mathcal{E}_\lambda + \mathcal{I}) \begin{pmatrix} v \\ \mu \end{pmatrix} = 0, \quad \begin{pmatrix} v \\ \mu \end{pmatrix} \in V^2_\alpha. \tag{9.61}$$

By (9.60) we have

$$\begin{pmatrix} v \\ \mu \end{pmatrix} = -\mathcal{E}_\lambda \begin{pmatrix} v \\ \mu \end{pmatrix} \in V^\infty_\alpha.$$

Hence, by invertibility of $\mathcal{E}_\lambda + \mathcal{I}$ in V^∞_α, $\begin{pmatrix} v \\ \mu \end{pmatrix} = 0$ is the only solution of (9.61). By a known property of compact operators (see e.g. Theorem 2, Chap. XIII, Sec. 1 in [28]), the uniqueness of the solution to the homogeneous equation (9.61) implies that the operator $\mathcal{E}_\lambda + \mathcal{I}$ is invertible in V^2_α. In turn, this implies that the operator (9.59) is invertible in V^2_α for any $\lambda \in (-(2kl)^{-1}, (2kl)^{-1})$. Hence equation (9.56) has a unique solution for any $(\chi^3, \chi^7) \in V^2_\alpha$ and

$$\left\| \begin{matrix} v \\ \mu \end{matrix} \right\|_2 \le 2kl \left\| \begin{matrix} \chi^3 \\ \chi^7 \end{matrix} \right\|_2. \tag{9.62}$$

Choose now $\chi^3 = 0$. Using definition (3.29), from (9.56) we obtain

$$\mathcal{Q}^{\bar h}_\alpha v = \chi^7.$$

In view of (9.62), we obtain (9.55). ■

PROPOSITION 9.8. *If* (H) *holds, then there exist* $\alpha, \gamma > 0$ *such that* (A3) *is satisfied.*

Proof. By (9.55) we find that $(-(2kl)^{-1},(2kl)^{-1})\not\subset\sigma$, where σ is the spectrum of $\mathcal{Q}^{\bar h}_\alpha$. On the other hand, by (9.48), $\sigma\subset[0,\infty]$. Hence (9.47) is satisfied. This implies (9.45), which is equivalent to

$$(v,\mathcal{Q}^{\bar h}_\alpha v)\ge(2kl)^{-1}\|\varGamma^{\bar h}_\alpha v\|_2^2\quad\text{for all } v\in L^2(0,1;\mathbb{R}^m). \tag{9.63}$$

Since

$$\|\varGamma^{\bar h}_\alpha-\varGamma_\alpha\|_{L^2\to L^2}\to 0\quad\text{and}\quad\|\mathcal{Q}^{\bar h}_\alpha-\mathcal{Q}_\alpha\|_{L^2\to L^2}\to 0\quad\text{as }\alpha\to 0,$$

(9.63) implies that there exists $\alpha>0$ such that

$$(v,\mathcal{Q}_\alpha v)\ge(4kl)^{-1}\|\varGamma_\alpha v\|_2\quad\text{for all } v\in L^2(0,1;\mathbb{R}^m),$$

i.e., (3.28) holds with $\gamma=(4kl)^{-1}$. ■

Bibliographical note. The idea of the proof of necessity of the conditions of Lipschitz stability, based on local equivalence to problems with equality constraints, was introduced in [16]. In that paper the special form (9.2) of the space of parameters was considered. Sensitivity of solutions was not discussed there. Instead, an abstract theorem due to Dontchev [10] was used in the proof of necessity.

10. Conclusion

Combining Theorems 7.5 and 8.1 with Propositions 9.4 and 9.8, we obtain the following theorem, which is the principal result of this paper.

THEOREM 10.1. *Suppose that* (I)–(III) *hold. If*

(i) (A1)–(A3) *are satisfied*

then

(ii) *there exist constants* $\sigma_1,\sigma_2,l>0$ *such that for each* $h\in\mathcal{B}^H_{\sigma_1}(h_0)$ *there is a unique solution* (x_h,u_h) *in* $\mathcal{B}^{X^\infty}_{\sigma_2}((x_0,u_0))$ *of* $(\mathrm{O})_h$ *and a unique associated Lagrange multiplier* $(p_h,\varrho_h,\nu_h)\in Y^\infty$. *Moreover*

$$\|x_{h_2}-x_{h_1}\|_{1,\infty},\|u_{h_2}-u_{h_1}\|_\infty,\|p_{h_2}-p_{h_1}\|_{1,\infty},|\varrho_{h_2}-\varrho_{h_1}|,\|\nu_{h_2}-\nu_{h_1}\|_\infty \le l\|h_2-h_1\|_H\quad\textit{for all } h_1,h_2\in\mathcal{B}^H_{\sigma_1}(h_0).$$

The mappings

$$(x_h,u_h):\mathcal{B}^H_{\sigma_1}(h_0)\to X^s,\quad (p_h,\varrho_h,\nu_h):\mathcal{B}^H_{\sigma_1}(h_0)\to Y^s$$

are Bouligand differentiable functions for any $s\in[2,\infty)$. *The B-differentials at* h_0 *in a given direction* $g\in H$ *are given by the solution and Lagrange multiplier of the linear-quadratic problem* $(\mathrm{L})_{h_0,g}$.

If in addition the strong dependence condition (A4) *holds*, *then* (ii) *implies* (i).

Theorem 10.1 provides a full characterization of Lipschitz stability and Bouligand differentiability results for $(\mathrm{O})_h$, in the sense formulated in Section 4. Example $(\mathrm{E})_h$ shows that we cannot expect differentiability of the solutions in X^∞. On the other hand, some weakening of the assumptions could be possible, if strong dependence condition (A4) is not satisfied, or if we are interested in differentiability properties of the solutions (or ε-solutions) only in a given direction, as in [50]. Also, it seems that some weakening of the constraint qualifications can be expected, if we are interested in stability and sensitivity of the solutions, but not of Lagrange multipliers.

On the basis of the methodology presented here, convergence analysis for approximations to optimal control problems, subject to mixed control-state constraints, can be performed. For Euler approximation, such an analysis is presented in [38]. Also the same methodology can be used to obtain a local convergence rate for Lagrange–Newton type optimization algorithms (see, e.g., [4]), including the mesh-independence principle for discretized problems [3].

Thus, the stability and sensitivity analysis for smooth nonlinear optimal control problems with mixed control-state constraints is fairly complete. Such an analysis is much less developed for optimal control problems if pure state constraints are present. In this case, the difficulties connected with the two-norm discrepancy are much harder to overcome. Moreover, pointwise variations of all constraints are no longer possible. For problems with first-order state constraints, sufficient conditions of Lipschitz continuity in X^2 as well as of directional differentiability have been obtained (see [35, 14]), but it is not known if these conditions are also necessary. Similarly, the problem of B-differentiability still remains open.

Acknowledgments. The author would like to express his gratitude to Maïtine Bergounioux and Helmut Maurer for careful reading of the manuscript, detecting numerous mistakes and suggesting some improvements.

References

[1] W. Alt, *Parametric programming with applications to optimal control and sequential quadratic programming*, Bayreuth. Math. Schr. 34 (1990), 1–37.

[2] —, *Stability of Solutions and the Lagrange–Newton Method for Nonlinear Optimization and Optimal Control Problems*, Habilitationsschrift, Univ. Bayreuth, Bayreuth, 1990.

[3] —, *Discretization and mesh-independence of Newton's method for generalized equations*, in: Mathematical Programming with Data Perturbations, A. V. Fiacco (ed.), Dekker, New York, 1998, 1–30.

[4] W. Alt and K. Malanowski, *The Lagrange–Newton method for nonlinear optimal control problems*, Comput. Optim. Appl. 2 (1993), 77–100.

[5] J. F. Bonnans and R. Cominetti, *Perturbed optimization in Banach spaces*, *III*: *Semi-infinite optimization*, SIAM J. Control Optim. 34 (1996), 1555–1567.

[6] J. F. Bonnans, R. Cominetti and A. Shapiro, *Sensitivity analysis of optimization problems under abstract constraints*, Math. Oper. Res. 23 (1998), 806–831.

[7] J. F. Bonnans and A. Shapiro, *Optimization problems with perturbations*, *A guided tour*, SIAM Rev. 40 (1998), 202–227.

[8] —, —, *Perturbation Analysis of Optimization Problem*, Springer, New York, 2000.

[9] A. L. Dontchev, *Implicit function theorems for generalized equations*, Math. Program. 70 (1995), 91–106.

[10] —, *Characterization of Lipschitz stability in optimization*, in: Recent Developments in Well-Posed Variational Problems, R. Lucchetti and J. Revalski (eds.), Kluwer, 1995, 95–116.

[11] A. L. Dontchev, *The Graves theorem revisited*, J. Convex Anal. 3 (1996), 45–53.

[12] —, *A proof of the necessity of linear independence condition and strong second-order sufficient optimality condition for Lipschitzian stability in nonlinear programming*, J. Optim. Theory Appl. 98 (1998), 467–473.

[13] A. L. Dontchev and W. W. Hager, *Lipschitz stability in nonlinear control and optimization*, SIAM J. Control Optim. 31 (1993), 569–603.

[14] —, —, *Lipschitzian stability for state constrained nonlinear optimal control*, ibid. 36 (1998), 698–718.

[15] A. L. Dontchev, W. W. Hager, A. B. Poore and B. Yang, *Optimality, stability and convergence in nonlinear control*, Appl. Math. Optim. 31 (1995), 297–326.

[16] A. L. Dontchev and K. Malanowski, *A characterization of Lipschitzian stability in optimal control*, in: Calculus of Variations and Optimal Control, A. Ioffe, S. Reich and I. Shafrir (eds.), CRC Res. Notes in Math. 411, Chapman & Hall, Boca Raton, 1999, 62–76.

[17] A. L. Dontchev and R. T. Rockafellar, *Characterizations of Lipschitz stability in nonlinear programming*, in: Mathematical Programming with Data Perturbations, A. V. Fiacco (ed.), Dekker, New York, 1998, 65–82.

[18] A. V. Fiacco, *Sensitivity analysis for nonlinear programming using penalty methods*, Math. Programming 10 (1976), 287–311.

[19] W. W. Hager, *Lipschitz continuity for constrained processes*, SIAM J. Control Optim. 17 (1979), 321–337.

[20] W. W. Hager and S. K. Mitter, *Lagrange duality theory for convex control problems*, ibid. 14 (1976), 843–856.

[21] A. Haraux, *How to differentiate the projection on a convex set in Hilbert space. Some applications to variational inequalities*, J. Math. Soc. Japan 29 (1977), 615–631.

[22] M. R. Hestenes, *Calculus of Variations and Optimal Control Theory*, Wiley, New York, 1966.

[23] A. D. Ioffe, *Necessary and sufficient conditions for a local minimum. 3: Second order conditions and augmented duality*, SIAM J. Control Optim. 17 (1979), 266–288.

[24] A. D. Ioffe and V. M. Tikhomirov, *Theory of Extremal Problems*, Nauka, Moscow, 1974 (in Russian); English transl.: North-Holland, Amsterdam, 1979.

[25] K. Ito and K. Kunisch, *Sensitivity analysis of solutions to optimization problems in Hilbert spaces with applications to optimal control and estimation*, J. Differential Equations 99 (1992), 1–40.

[26] K. Jittorntrum, *Sequential Algorithms in Nonlinear Programming*, Ph.D. thesis, Australian National Univ., Canberra, 1978.

[27] —, *Solution point differentiability without strict complementarity in nonlinear programming*, Math. Programming Study 21 (1984), 127–138.

[28] L. V. Kantorovich and G. P. Akilov, *Functional Analysis*, Nauka, Moscow, 1977 (in Russian); English ed.: Pergamon Press, Oxford, 1982.

[29] D. Klatte, *On regularity and stability in semi-infinite optimization*, Set-Valued Anal. 3 (1995), 101–111.

[30] D. Klatte and R. Henrion, *Regularity and stability in nonlinear semi-infinite optimization*, Nonconvex Optim. Appl. 25 (1997), 69–102.

[31] D. G. Luenberger, *Optimization by Vector Space Methods*, Wiley, New York, 1974.

[32] K. Malanowski, *Differential stability of solutions to convex control constrained optimal control problems*, Appl. Math. Optim. 12 (1984), 1–14.

[33] —, *Sensitivity analysis of optimization problems in Hilbert space with application to optimal control*, ibid. 21 (1990), 1–20.

[34] —, *Second order conditions and constraint qualifications in stability and sensitivity analysis of solutions to optimization problems in Hilbert spaces*, ibid. 25 (1992), 51–79.

[35] K. Malanowski, *Two-norm approach in stability and sensitivity analysis of optimization and optimal control problems*, Adv. Math. Sci. Appl. 2 (1993), 397–443.
[36] —, *Stability and sensitivity of solutions to nonlinear optimal control problems*, Appl. Math. Optim. 32 (1995), 111–141.
[37] —, *Bouligand differentiability of solutions to parametric optimal control problems*, preprint WP-1-2000, Systems Research Institute, Polish Acad. Sci., Warszawa, 2000.
[38] K. Malanowski, C. Büskens and H. Maurer, *Convergence of approximations to nonlinear optimal control problems*, in: Mathematical Programming with Data Perturbations, A. V. Fiacco (ed.), Dekker, New York, 1998, 253–284.
[39] K. Malanowski and H. Maurer, *Sensitivity analysis for parametric optimal control problems with control-state constraints*, Comput. Optim. Appl. 5 (1996), 253–283.
[40] H. Maurer, *First- and second-order sufficient optimality conditions in mathematical programming and optimal control*, Math. Programmming Study 14 (1981), 163–177.
[41] H. Maurer and S. Pickenhain, *Second order sufficient conditions for optimal control problems with mixed control-state constraints*, J. Optim. Theory Appl. 86 (1995), 649–667.
[42] F. Mignot, *Contrôle dans les inéquations variationnelles*, J. Funct. Anal. 22 (1976), 130–185.
[43] S. M. Robinson, *Perturbed Kuhn–Tucker points and rate of convergence for a class of nonlinear programming algorithms*, Math. Programming 7 (1974), 1–16.
[44] —, *Strongly regular generalized equations*, Math. Oper. Res. 5 (1980), 43–62.
[45] —, *Local structure of feasible sets in nonlinear programming*, *Part III*: *Stability and sensitivity*, Math. Programming Study 30 (1987), 97–116.
[46] —, *An implicit-function theorem for a class of nonsmooth functions*, Math. Oper. Res. 16 (1991), 292–309.
[47] W. Rudin, *Functional Analysis*, McGraw-Hill, New York, 1991.
[48] A. Shapiro, *On concepts of directional differentiability*, J. Math. Anal. Appl. 66 (1990), 477–487.
[49] —, *Perturbation analysis of optimization problems in Banach spaces*, Numer. Funct. Anal. Optim. 13 (1992), 97–116.
[50] A. Shapiro and J. F. Bonnans, *Sensitivity analysis of parametrized programs under cone constraints*, SIAM J. Control Optim. 30 (1992), 97–116.
[51] J. Sokołowski, *Sensitivity analysis of control constrained optimal control problems for distributed parameter systems*, ibid. 26 (1987), 1542–1556.
[52] J. Sokołowski and J.-P. Zolesio, *Introduction to Shape Optimization. Shape Sensitivity Analysis*, Springer, Berlin, 1992.
[53] K. Yosida, *Functional Analysis*, Springer, Berlin, 1980.
[54] V. Zeidan, *The Riccati equation for optimal control problems with mixed control-state constraints*: *necessity and sufficiency*, SIAM J. Control Optim. 32 (1994), 1297–1321.